Terrorist Content
Analytics Platform

MAPPING FAR-RIGHT TERRORIST PROPAGANDA ONLINE

May 2024

Powered by
tech
against
terrorism

Launched in 2020, Tech Against Terrorism's Terrorist Content Analytics Platform (TCAP) is a secure online tool that detects and verifies terrorist content and then alerts technology companies to the presence of such material on their platforms. In the three and a half years since its launch, the TCAP has had an incredible impact on countering terrorist use of the internet, alerting 133 different tech platforms to over 30,000 pieces of terrorist content, of which 72% is now offline.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This report evaluates the distribution of far-right terrorist propaganda across online platforms between 20 February 2021 and 1 November 2023. The report draws on data from the Terrorist Content Analytics Platform (TCAP), the world's largest database of verified terrorist content utilised by tech platforms.

The report maps the evolution of the far-right terrorist threat online by analysing the type and affiliation of propaganda that is most prevalent online and on which platforms it is found, as well as how tactics of exploitation have adapted to content moderation efforts. We evaluate tech platform responses to far-right terrorist propaganda, and we offer recommendations for improved mitigation strategies to the evolving threat landscape and exploitation of new technologies.

MAPPING FAR-RIGHT TERRORIST PROPAGANDA ONLINE

Far-right terrorist content is prevalent across a wide range of online services, and especially on encrypted messaging apps and alt-tech video-sharing platforms where its moderation by platforms is inconsistent.

We have identified that an increasing proportion of far-right terrorist content is found on mainstream social media platforms. This content is often edited or gamified to circumvent automated content moderation systems, which in some cases has allowed this material to be viewed a significant number of times before being removed.

The veneration of individual far-right terrorists, through the re-sharing of their manifestos and livestreams, was more common than the sharing of propaganda produced by far-right terrorist organisations.

Content produced by the perpetrator of the 2019 Christchurch attack has received unparalleled attention among far-right online networks we monitor, and its continued re-purposing underscores the risk that this material inspires similar violent attacks.

The significant volume of Atomwaffen Division (AWD) content that we identified highlights the organisation's potency as a militant accelerationist brand despite its formal dissolution.

KEY TRENDS

Sanctification: The veneration of far-right violent extremists by online supporter networks through the production of slick propaganda and edited versions of livestreams is a concerning trend that trivialises violence and can radicalise vulnerable online users.

Gamification: The gamification of real-world acts of far-right violence is a popular tactic used by online extremists to build communities across platforms, radicalise younger users, and incite imitative violence through its normalisation.

FUTURE RISKS

Tech Against Terrorism assesses that the early experimentation with Generative AI by far-right actors indicates an emerging threat of exploitation in the medium to long term. We highlight the following risks of exploitation of Generative AI for far-right visual propaganda:

Sanctification: Generative AI is likely to bolster the creation, by means of synthetic text or editing tools, of propaganda that glorifies terrorist actors, as well as enhance its dissemination.

Circumvention: Generative AI is highly likely to be used by terrorist and violent extremist (TVE) actors to circumvent existing content moderation systems, including hashing, through creative editing, media spawning and variant recycling.

Event-specific propaganda and deepfakes: Generative AI is likely to be widely used by extremists in the long-term future to exploit crises through the rapid creation and dissemination of propaganda or deepfakes that use disinformation to drive polarisation and incite violence.

Terrorist Content Analytics Platform

Powered by tech against terrorism

FAR-RIGHT TERRORIST PROPAGANDA IN NUMBERS

Between February 2021 and October 2023:

2,966 URLs

Containing far-right terrorist content were identified by Tech Against Terrorism.

We alerted 2,348 of these URLs to 55 different online platforms through the Terrorist Content Analytics Platform (TCAP).

130 different platforms

With far-right terrorist propaganda were identified by Tech Against Terrorism.

We identified continued exploitation of messaging and video- sharing platforms and increased targeting of mainstream social media platforms.

100 far-right URLs

Or more, are identified by Tech Against Terrorism every month, with the same platforms persistently targeted using similar methods of exploitation.

X4 Christchurch content

Between October 2022 and October 2023, the volume of content Tech Against Terrorism identified relating to the Christchurch attack perpetrator more than quadrupled from 242 URLs to 1098 URLs

15% Uncooperative platforms

Tech Against Terrorism was unable to alert 15% of far-right terrorist content identified due to its hosting on uncooperative tech platforms.

66% Lower removal rates

Around two thirds (66%) of far-right terrorist content alerted through the TCAP had been removed by tech companies (compared to 78% removal for alerted Islamist content).

MITIGATING THE FAR-RIGHT TERRORIST THREAT

HOW ARE TECH PLATFORMS RESPONDING?

Uncooperative platforms

Tech Against Terrorism was unable to alert 15% of far-right terrorist content identified due to it being hosted on uncooperative tech platforms. Smaller alt-tech video-sharing and social media platforms were particularly uncooperative despite being heavily exploited by extremists.

Lower removal rates

Around two thirds (66%) of far-right terrorist content alerted through the TCAP had been removed by tech companies. There are likely to be many reasons for this, which include a lack of clarity on the illegality of far-right content, jurisdictional gaps and confusion, and a lack of expertise among moderators.

Inconsistent content moderation

There was a wide disparity in removal rates across different platforms and types of terrorist content, and inconsistent enforcement of platform policies. This suggests a lack of clarity in the tech sector on what constitutes an adequate response to far-right terrorist content.

POLICY RECOMMENDATIONS FOR TECH PLATFORMS

UNDERSTANDING THE THREAT

⇒ We recommend that tech companies monitor for far-right terrorist-affiliated symbols and key words to improve detection of far-right terrorist content. Tech Against Terrorism's Knowledge Sharing Platform (KSP) provides an easily navigable and extensive database of far-right symbols and phrases for platform moderators. Access it here.

LEGAL CLARITY

⇒ We recommend that tech companies prohibit both terrorism and violent extremism in their Terms of Service / Community Guidelines.[1] A prohibition of terrorism should include content which encourages, supports, glorifies and/or promotes terrorism, terrorist organisations, terrorist attacks and attackers.

⇒ We recommend consulting national and supranational designation lists as a guide to the far-right entities that have been designated as terrorist through a stringent legal process.[2] Find our list of designated far-right groups here (and in Fig. 2).

⇒ We recommend that tech companies focus on content produced by far-right terrorist attack perpetrators, such as the manifesto and livestream produced by the Christchurch attacker. Find attacker-produced content included in the TCAP here.

TRANSPARENCY

⇒ We recommend that tech companies explain what is prohibited on their services in a way that is clear and easily understandable for users.

⇒ We recommend informing users why action has been taken against their content or account, and on what grounds, with reference to a specific policy violation.[3]

⇒ We recommend that tech platforms produce a transparency report about their moderation enforcement actions.

[1] The prohibition of violent extremism allows platforms to more effectively moderate violent far-right content that has not been produced by designated far-right groups.

[2] Canada, the UK, and Australia have the most developed designation lists for far-right entities. For more information on international designations systems and the implications for online terrorist content, see Tech Against Terrorism's report, 'Who Designated Terrorism? The Need for Legal Clarity to Moderate Terrorist Content Online'

[3] Tech platforms are encouraged to provide an explanation of why a user's content has been removed under the EU's Terrorist Content Online (TCO) Regulation and the Digital Services Act (DSA).

INTRODUCTION

This report presents the principal developments and trends in far-right terrorist exploitation of the internet since the Terrorist Content Analytics Platform (TCAP) began alerting far-right terrorist content in February 2021. The report explores the discernible patterns in the distribution of far-right terrorist content across online platforms, the trends emerging in the exploitation of these services, and the nature of tech platform responses.

Proponents of far-right militant accelerationism[4] are consistently and strategically exploiting online platforms to share propaganda that glorifies and incites violence, sanctifies terrorists and violent extremists, and gamifies terrorist attacks. It is a realistic possibility that this potent accelerationist brand that trivialises violence and has already inspired offline attacks, will contribute to the radicalisation of vulnerable online users in the future.[5]

DEFINING FAR-RIGHT TERRORIST CONTENT

The purpose of this report is not to engage with definitional debates on what constitutes the far-right. However, given that we are referring to 'far-right terrorist content' throughout the report it is necessary to set out how we have defined this content for the purpose of collecting and analysing data.

Below, we outline the scope of far-right content included within this report as determined by the TCAP's Tiered Inclusion Policy.[6] You can find a more detailed explanation of our Inclusion Policy in the Annex.

TIER 2 – CRISIS

We include the crisis material (manifestos and/or livestream) produced by the far-right perpetrators of the following attacks:

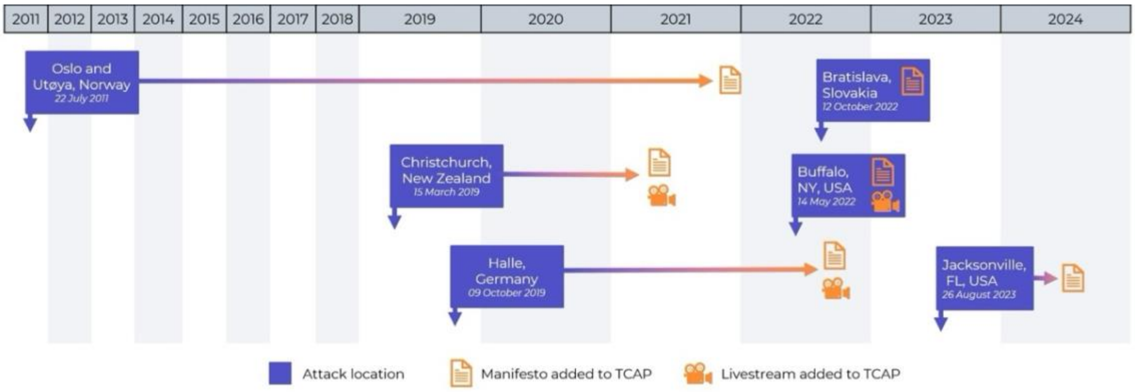

2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
Oslo and Utøya, Norway 22 July 2011
Bratislava, Slovakia 12 October 2022
Christchurch, New Zealand 15 March 2019
Buffalo, NY, USA 14 May 2022
Halle, Germany 09 October 2019
Jacksonville, FL, USA 26 August 2023
Attack location
Manifesto added to TCAP
Livestream added to TCAP

Figure 1: Far-right terrorist incidents in scope of the TCAP Inclusion Policy.

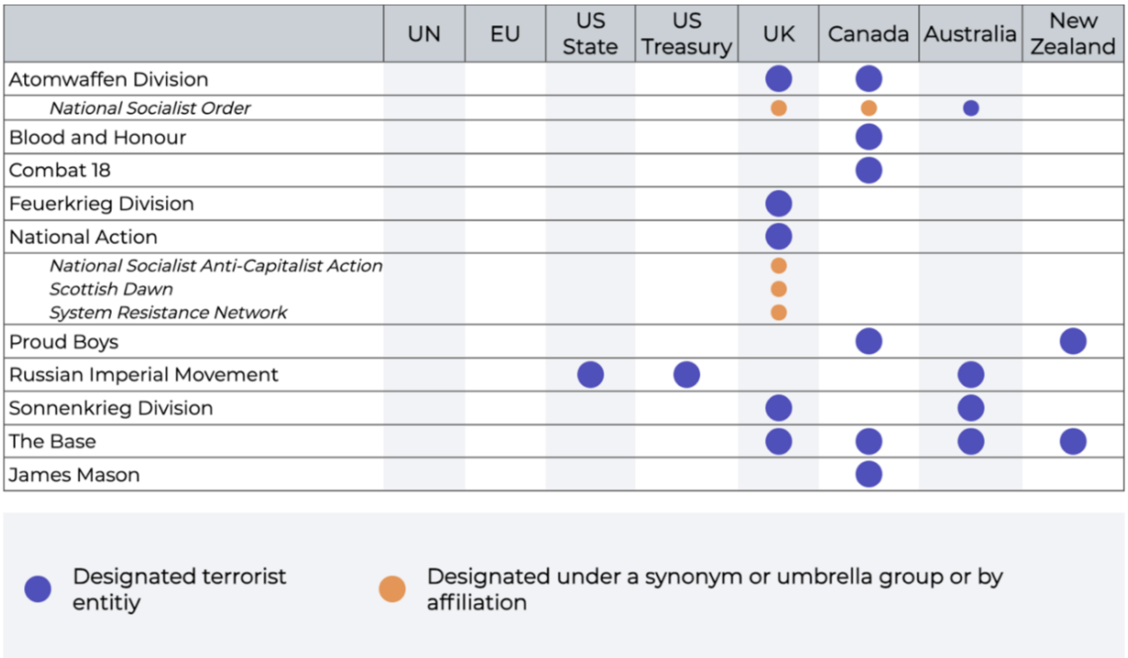
[4] Far-right accelerationism holds that acts of mass violence will hasten the collapse of socio-political systems believed to be systematically oppressing white people and will eventually lead to the establishment of a desired white ethnostate. Source: https://www.rsis.edu.sg/rsis-publication/icpvtr/bratislava-shooting-the-making-of-terrorgrams-first-saint/

[5] The perpetrator of an attack in Bratislava, Slovakia, that killed two members of the LGBTQ+ community was heavily influenced by the online militant accelerationist community known as "Terrorgram."

[6] You can find our full Inclusion Policy on the TCAP website here or in the Annex to this report.

TIER 3 – DESIGNATION

We include the official propaganda produced by the following 14 far-right terrorist entities, which are designated as terrorist by the authorities listed:



	UN	EU	US State	US Treasury	UK	Canada	Australia	New Zealand
Atomwaffen Division					🔵	🔵		
National Socialist Order					🟠	🟠	🔵	
Blood and Honour						🔵		
Combat 18						🔵		
Feuerkrieg Division					🔵			
National Action					🔵			
National Socialist Anti-Capitalist Action					🟠			
Scottish Dawn					🟠			
System Resistance Network					🟠			
Proud Boys						🔵		🔵
Russian Imperial Movement			🔵	🔵			🔵	
Sonnenkrieg Division					🔵		🔵	
The Base					🔵	🔵	🔵	🔵
James Mason						🔵		

🔵 Designated terrorist entitiy

🟠 Designated under a synonym or umbrella group or by affiliation

Figure 2: Far-right designated terrorist entities for TCAP Inclusion Policy.

TIER 4 – PROMOTIONAL

INSPIRATIONAL MATERIAL[7]

During our data collection period, the following far-right inspirational material was in scope:

- Gamified versions of the Christchurch attack livestream
- Gamified versions of the Buffalo attack livestream
- Gamified versions of the Halle attack livestream

[7] Inspirational material is a sub-tier of Tier 4 that includes "content that explicitly encourages, glorifies and/or incites a terrorist act or praises the perpetrator(s) of that act, given the entity (individual or organisation) is included within scope of the TCAP."

1. TERRORIST CONTENT ANALYTICS PLATFORM (TCAP) DATA ANALYSIS

KEY FINDINGS

PATTERNS OF FAR-RIGHT TERRORIST PROPAGANDA DISSEMINATION

Between February 2021 and October 2023, Tech Against Terrorism identified **2,966 URLs** containing far-right terrorist content on **130** different online platforms.

In the same period, we alerted **2,348** of these URLs to **55** different online platforms through the TCAP, **66%** of which were removed.

There has been a steady improvement in Tech Against Terrorism's ability to identify and alert far-right terrorist content online over time, with at least **100 URLs** now consistently identified every month.

Tech Against Terrorism has consistently identified that a higher volume of terrorist content relates to **'lone-actor'** terrorists than to designated far-right **groups**, which suggests that attacker-produced content is more prominent in the far-right spaces we monitor.

Christchurch attack content is by far the type of terrorist content identified most prominently and consistently in the far-right online networks monitored by Tech Against Terrorism, with **1,098** TCAP submissions. Between October 2022 and October 2023, the number of TCAP submissions relating to Christchurch content more than **quadrupled**, in large part due to it being increasingly discoverable on mainstream social media platforms. Within the Christchurch content identified on such platforms, a significant proportion comprises edited or gamified versions of the livestreamed attack.

Atomwaffen Division (AWD) content, identifiable by branding, was by far the most widely disseminated of any far-right terrorist group across the online spaces monitored by Tech Against Terrorism, with **382 URLs** identified.

TECH PLATFORM RESPONSES

Smaller alt-tech **video-sharing** and **social media** platforms are being heavily exploited by extremists to host far-right terrorist content. Tech Against Terrorism is unable to engage with these platforms to alert them to this content. We were therefore unable to alert **15%** of the far-right content which we identified, comprising **448 URLs** on **84** different platforms.

As of 31 October 2023, around two thirds (**66%**) of far-right terrorist content alerted through the TCAP had been removed by tech companies (compared to **78%** removal of alerted Islamist content). There are many likely reasons for this, which include a lack of clarity on the illegality of far-right content, jurisdictional gaps and confusion, and a lack of expertise among moderators.

The rates at which content produced by different far-right terrorist organisations and entities is removed by platforms diverge greatly (between **100%** and **38%**), which further highlights the lack of clarity and consistency in the tech sector's response to far-right terrorist content.

PART 1: PATTERNS OF FAR-RIGHT TERRORIST PROPAGANDA DISSEMINATION

Overview of far-right terrorist content through key TCAP metrics

Metric	Description	Total
TCAP submissions	The number of unique URLs containing terrorist content submitted to the TCAP.	**2,966**
Alerts sent to tech platforms	The number of automated alerts sent to tech companies notifying them of terrorist content on their platform. Alerts are only sent to tech companies registered for TCAP alerts.	**2,348**
Percentage of alerted URLs offline	The percentage of content alerted to tech companies which is no longer accessible.	**66%**
Tech platforms alerted	The total number of tech platforms to which the TCAP has sent automated alerts.	**55**

The overall trend in far-right terrorist content over time

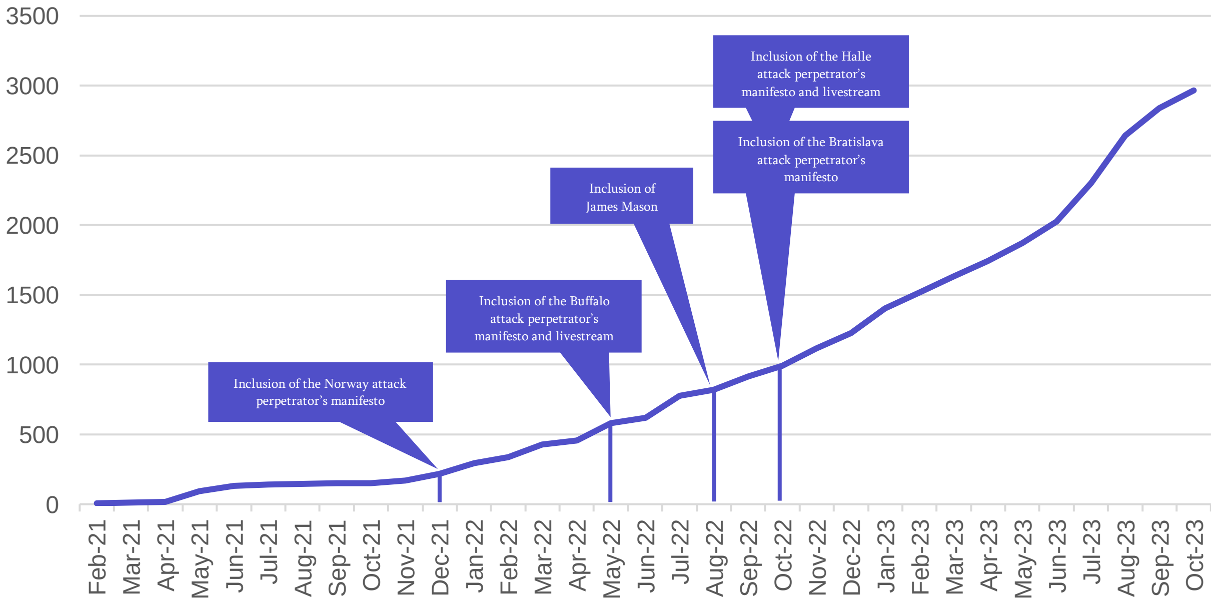


Figure 3: Cumulative TCAP submissions of far-right content over time.

Key finding

A steady improvement in Tech Against Terrorism's ability to identify and alert far-right terrorist content online.

The volume of far-right terrorist content submitted via the TCAP each month slowly increased during the first year of collection but remained inconsistent and below 100 URLs per month.

Since March 2022, there has been a steady increase in the volume of far-right terrorist content submitted per month, reaching a peak of 343 URLs submitted in August 2023. This upward trend suggests a consistent improvement in Tech Against Terrorism's ability to identify far-right terrorist content online.

However, the data does not necessarily provide evidence for a material increase in far-right terrorist content on the internet. The monthly submissions and alerts of far-right terrorist content via the TCAP are influenced by both external factors (such as the wider threat landscape) and internal factors (such as fluctuations in monitoring capacity).

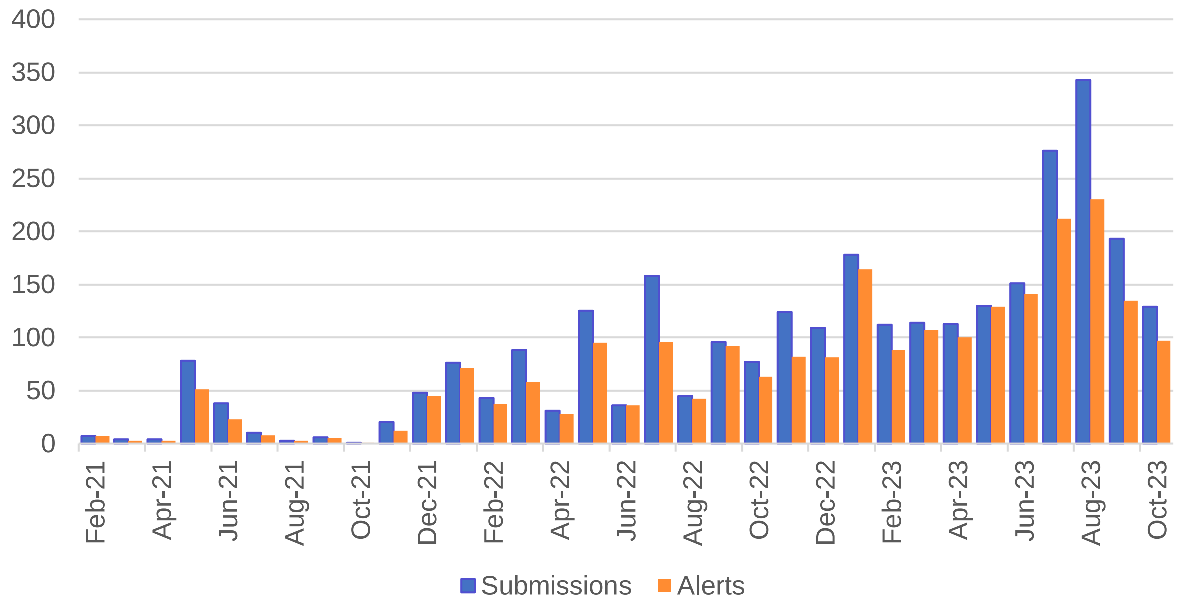


Figure 4: Far-right TCAP submissions and alerts over time (per month).

It is likely that multiple factors contribute to the variation in the volume of far-right terrorist content identified and submitted to the TCAP each month. These factors include but are not limited to:

- **The prevalence of far-right terrorist content on the online platforms monitored by Tech Against Terrorism.** In contrast with Islamist terrorist content, most of the far-right content we submit to the TCAP each month is not newly disseminated propaganda but reposted old content (e.g., historic manifestos or propaganda videos). Dissemination of this content is often supporter-driven and dictated by far-right responses to events such as violent attacks. Our proactive daily monitoring of known far-right channels means that any notable increase in output is likely to be reflected in increased TCAP submissions and alerts.

- **Changes in tech platform content moderation.** The availability of far-right terrorist content on online platforms correlates directly with adequacy of platforms' moderation policies and enforcement practices. Significant shifts in content moderation practices on platforms targeted by violent far-right actors can increase or reduce the volume of far-right terrorist content identified and alerted through the TCAP.

- **Expansion of TCAP's Inclusion Policy.** In February 2021, the TCAP Inclusion Policy only covered content produced by 13 designated far-right terrorist entities, as well as the Christchurch perpetrator. Since then, Tech Against Terrorism has added another designated far-right entity and five more attack perpetrators.[8] Each addition increased the range and volume of far-right terrorist content in scope for alerting.

- **Tech Against Terrorism's Open-Source Intelligence (OSINT) focus.** Since content collection for the TCAP is reliant on manual identification and verification, the OSINT team's operational balance between far-right monitoring and collection and other terrorist content therefore impacts TCAP submissions. The necessarily proactive nature of far-right terrorist content collection, and the greater allocation of operational effort it requires, makes this factor especially influential.

While impossible to quantify the relative contribution of each individual factor to the upward trend in far-right terrorist content, Tech Against Terrorism consistently identified and alerted a higher volume of far-right terrorist content between May and October 2023 than previously.

THE TCAP TIERED SYSTEM

Since the introduction of the TCAP Tiered System in July 2023, content related to far-right terrorism has been split across three tiers: crisis, designation, and promotional. This has allowed us to track more accurately the type of far-right terrorist content identified each month.

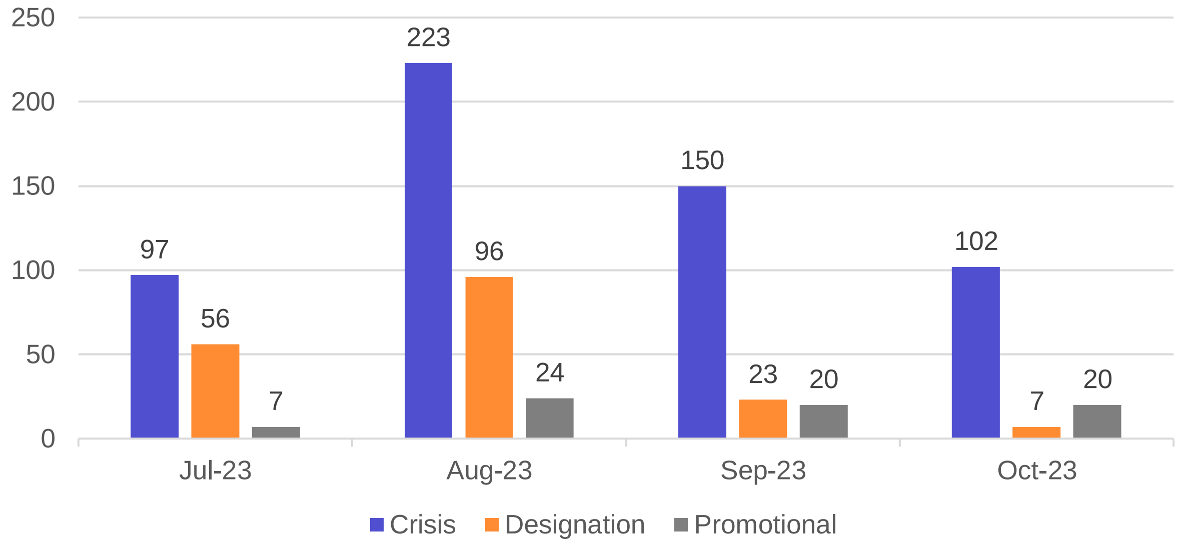


Figure 5: Far-right submissions by Tiered category of content since introduction of Tiered System in July 2023.

[8] See Fig. 1 and Fig. 3

Key finding

Tech Against Terrorism has consistently identified a higher volume of terrorist content relating to 'lone-actor' terrorists than to designated far-right groups, which suggests that attacker-produced content is more prominent in the far-right spaces we monitor.

Content produced by individual far-right terrorists (livestreams and manifestos) was identified more frequently between July and October 2023 than content produced by the designated terrorist groups in scope of the TCAP (see Fig. 5).

This is notable since the official content affiliated with designated organisations (such as videos, manuals, posters) vastly exceeds in variety and availability the content relating to individual 'lone actors', which comprises a small number of livestreams and manifestos which are in scope.

This suggests that far-right attacker-produced content plays an important role within the violent far-right online spaces which we monitor – more so than any one far-right group. Promotional far-right terrorist content (currently limited to gamified livestreams) represents a consistent but small proportion of monthly submissions.

GROUP VERSUS INDIVIDUAL-PRODUCED CONTENT

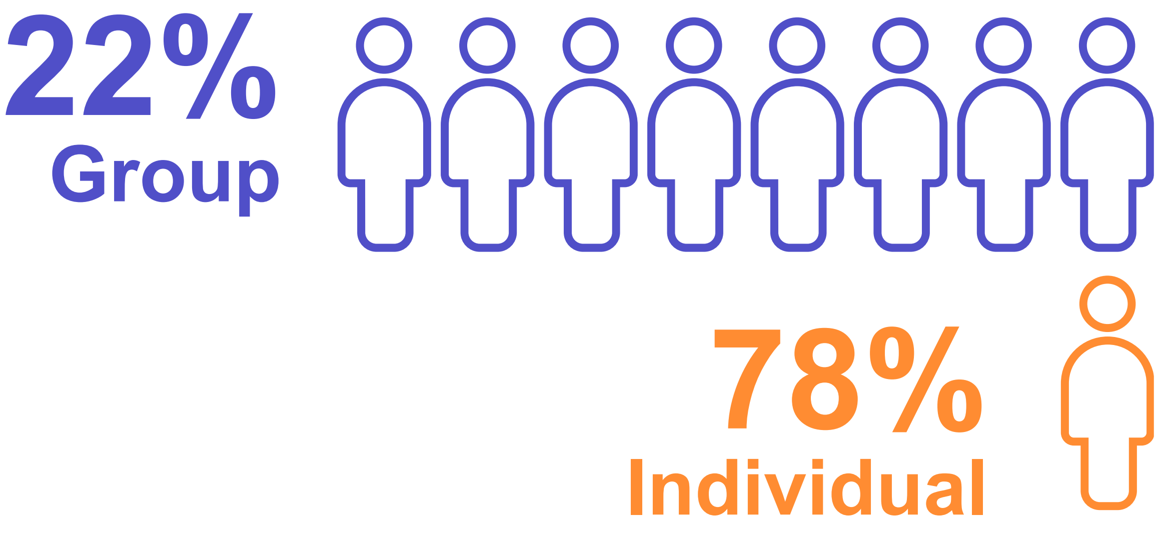


Figure 6: Percentage of far-right TCAP submissions relating to an individual versus a group.

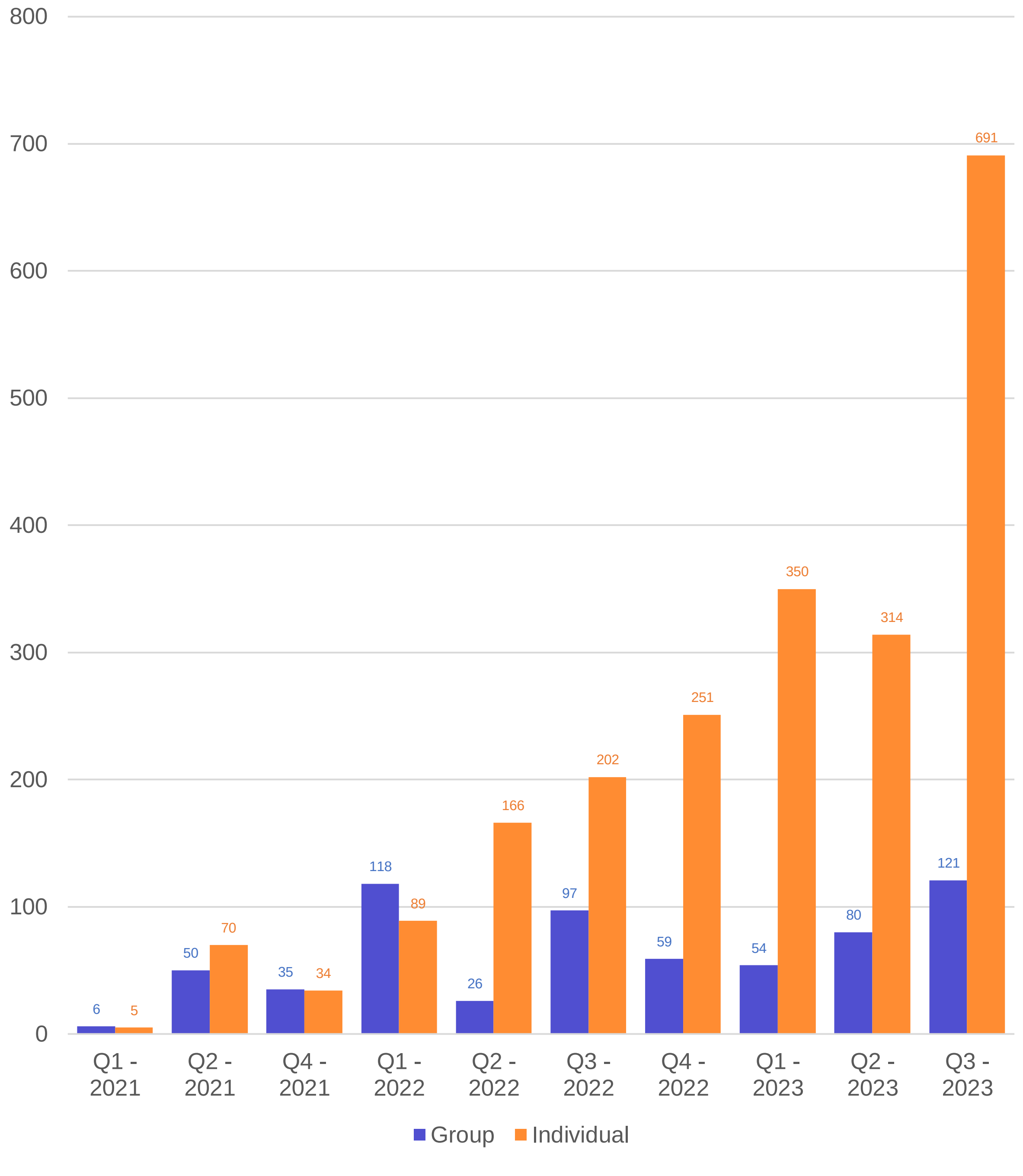


Figure 7: Far-right TCAP submissions relating to an individual versus a group by quarter.

Overall, almost 80% of far-right submissions relate to terrorist content produced by individuals. Since the second quarter of 2022, terrorist content identifiably produced by individuals has greatly exceeded in each quarter the content produced by groups.

The multiple factors which are likely to have contributed to this growing imbalance include: the addition of far-right individuals to the TCAP over time; organisational fractures within the far-right groups we monitor; and a continued shift away from organised groups to diffuse violent far-right online networks.

ANALYSIS: DISTRIBUTION OF FAR-RIGHT CONTENT BY TERRORIST ENTITY

Key finding

Christchurch attack content is by far the most prominent type of terrorist content consistently identified in the far-right online networks monitored by Tech Against Terrorism; however, Atomwaffen Division (AWD)-branded propaganda remains popular within these networks.

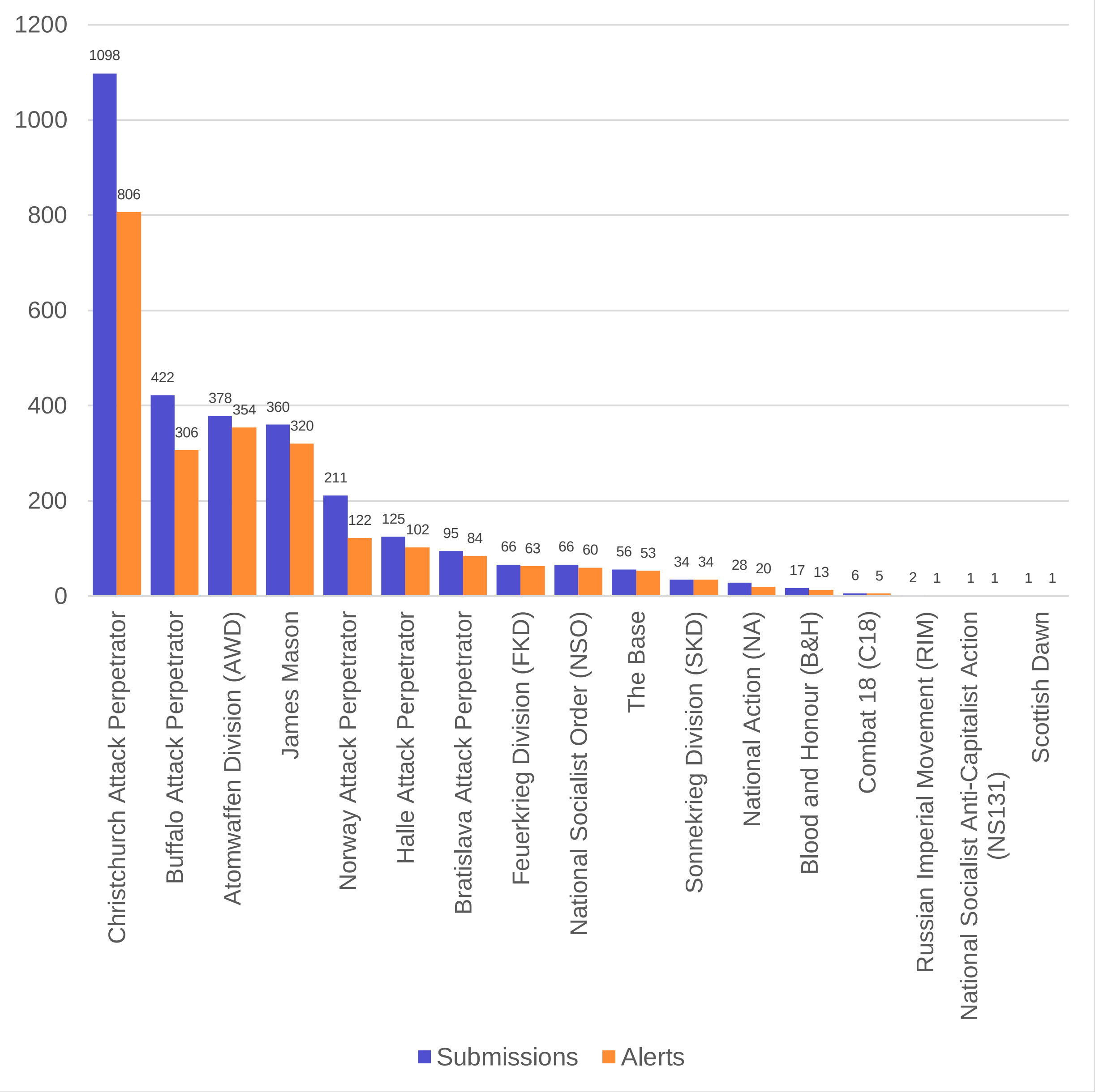


Figure 8: TCAP submissions and alerts by far-right terrorist entity.

Content produced by the perpetrator of the 2019 Christchurch attack has been submitted and alerted through the TCAP more often than any other piece of far-right terrorist content. In fact, submissions relating to Christchurch content were, at 1098 submissions, almost treble the volume for the second-most submitted content (Buffalo content at 424 submissions).

A significant volume of content produced by other far-right attackers has been identified, including the perpetrators of the attacks in Norway (210 submissions), Halle (102 submissions), and Bratislava (92 submissions). The variations by attacker in the volume of this material are determined primarily by the status (and veneration) of these actors in violent extremist networks but also by whether they produced a manifesto **and** livestreamed their attack (as is the case with Christchurch, Buffalo, and Halle), and how long they have been in scope of the TCAP.

James Mason is a unique entity: he is included in TCAP due to his designation as a terrorist ideologue rather than a perpetrator of terrorism. The volume of James Mason content we have identified (373 submissions) indicates the important role Mason plays as a propagandist for the violent far right but is also due to his output comprising a higher volume of content that is eligible for inclusion (including books, interviews, and podcasts) than other entities in scope.

Meanwhile, the volume of propaganda content which identifiably relates to designated far-right groups is relatively low. This is likely to be due to the ephemeral and decentralised nature of these groups, which undermines their ability to maintain consistent propaganda output. Internal fractures have led to groups becoming defunct and producing offshoots with new names and brands. One notable exception is Atomwaffen Division (382 submissions), which has been effective at establishing an international brand of violent propaganda that is reinforced by supporter-generated content and allows it to outlive the now-fractured group.

FAR-RIGHT CONTENT PRODUCED BY TERRORIST ATTACK PERPETRATORS

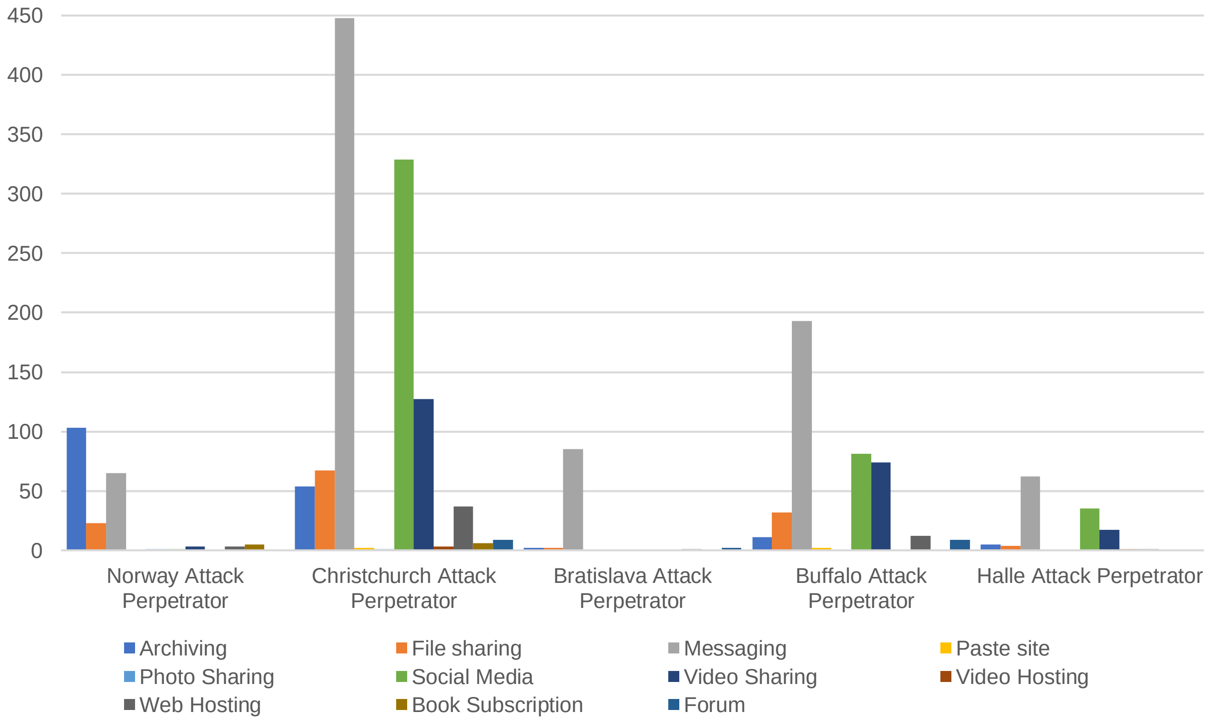


Figure 9: Submissions by far-right terrorist entity split by platform type.

A consistently high volume of far-right content is found on messaging platforms. The popularity of these platforms as a forum for interacting with far-right violent content is likely to be due to the stable environment they offer. The one notable exception to this is content produced by the Norway 2011 attacker, whose content is most often identified on archiving services. This is most likely to be due to the attack being by far the least recent (over ten years ago), such that it has less currency and salience for the users of messaging platforms.

Whereas Christchurch and Buffalo content is spread across many platform types (7+), Norway, Halle and Bratislava content is more heavily concentrated on fewer platform types. Bratislava content has only been found on one platform type (messaging).

This allocation suggests that Christchurch and Buffalo content resonates with a wider range of far-right networks and has achieved greater virality across the internet, which is almost certainly due to the availability and subsequent proliferation of an attacker livestream. We have identified an increasing volume of these livestreams (normally edited versions to avoid moderation) on mainstream social media platforms (see Fig. 10).

TCAP SUBMISSIONS RELATING TO FAR-RIGHT ATTACKERS IDENTIFIED ON SOCIAL MEDIA OVER TIME

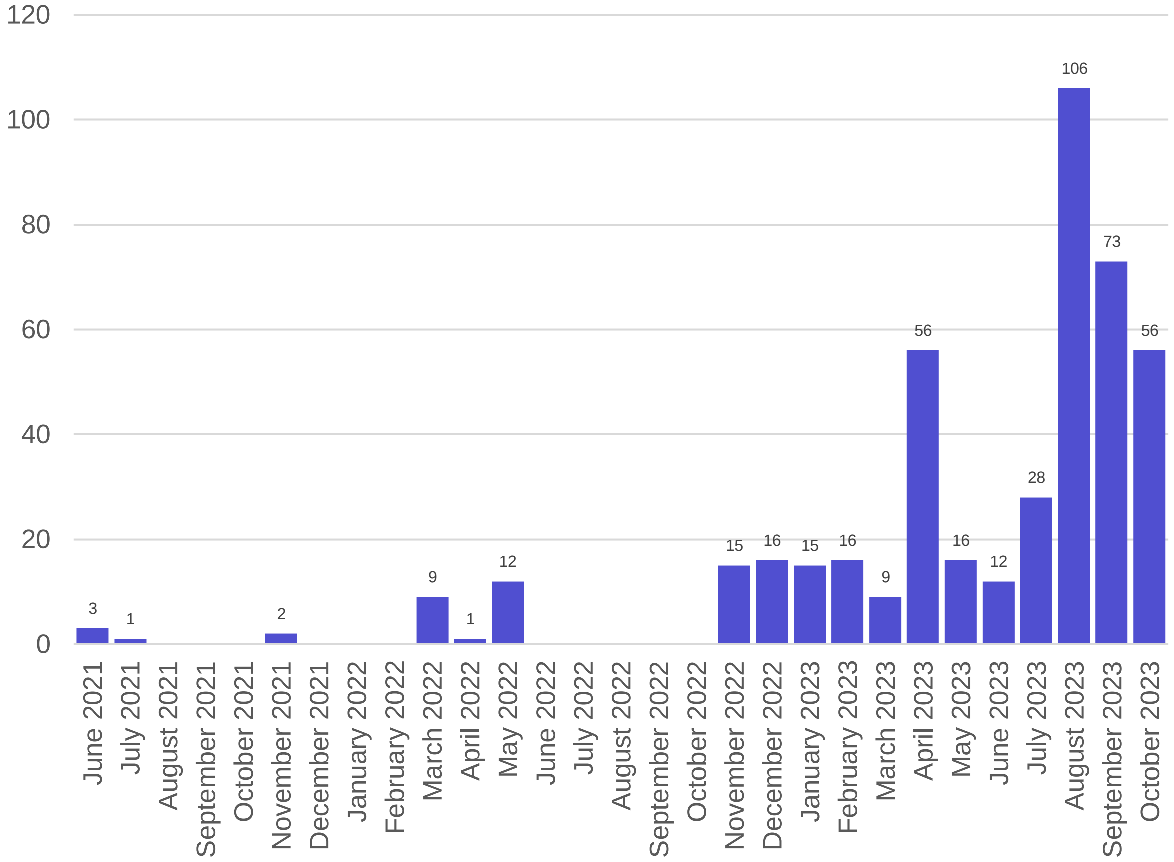


Figure 10: TCAP submissions relating to far-right attackers identified on social media platforms over time.

CASE STUDY: CHRISTCHURCH CONTENT

Key finding

Between October 2022 and October 2023, the number of TCAP submissions relating to Christchurch content more than quadrupled, in large part due to its increased discovery across mainstream social media platforms.

CUMULATIVE TCAP SUBMISSIONS OF CHRISTCHURCH CONTENT

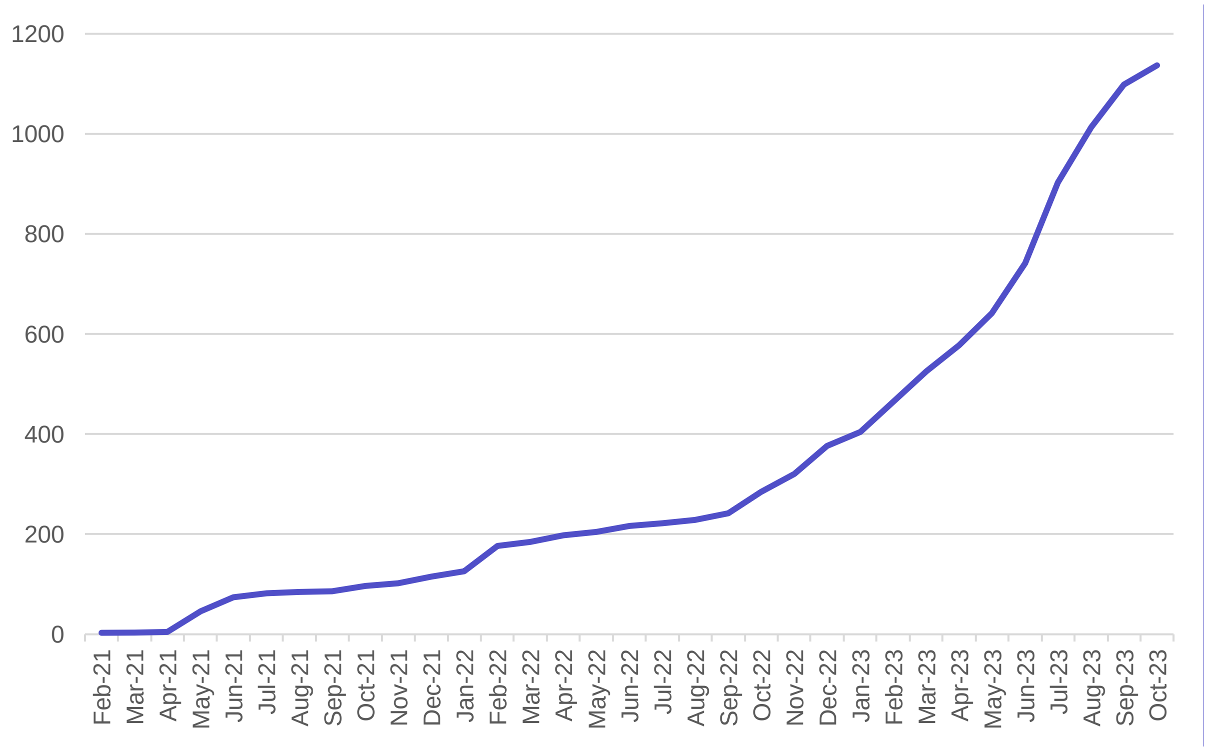


Figure 11: Cumulative TCAP submissions of Christchurch content.

There has been a significant and sustained increase over the past year in the identification and submission of URLs relating to Christchurch content. Between October 2022 and October 2023, the number of Christchurch submissions more than quadrupled from 242 URLs to 1098 URLs. This is largely due to its continued identification on messaging platforms alongside a huge increase in its identification on social media platforms. By comparison with 2022, there was a twelve-fold increase in 2023 in the Christchurch submissions identified on social media platforms in 2023 (26 URLs to 310 URLs).

This trend can partly be explained by an increased strategic focus on these platforms by our OSINT capabilities, but it nonetheless highlights the concerning proliferation of far-right terrorist content on mainstream social media platforms.

A significant proportion of the Christchurch content identified on social media platforms comprises edited or gamified versions of the Christchurch livestream. This demonstrates the ease with which content editing techniques can be used to circumvent the automated detection systems employed by large platforms. It also suggests that supporter networks of the violent far right are continuing to experiment with content moderation avoidance and that they remained determined to operate on mainstream platforms and disseminate propaganda to a larger audience.

CHRISTCHURCH SUBMISSIONS BY PLATFORM TYPE OVER TIME

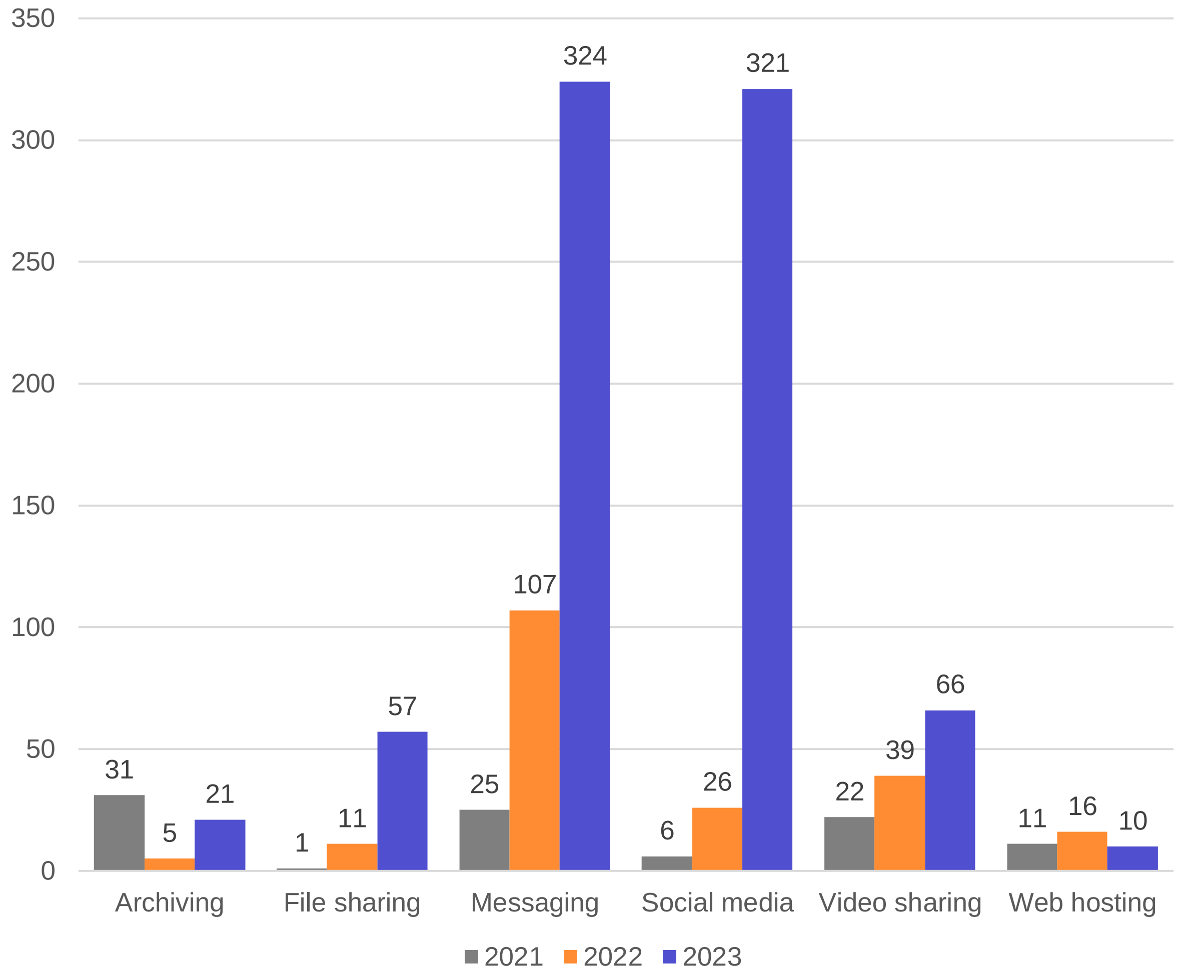


Figure 12: Christchurch submissions by selected platform type per calendar year.

A relatively small volume of Christchurch content was identified on video-sharing platforms (137 URLs), primarily consisting of smaller alt-tech platforms. These platforms have less stringent content moderation policies due to more 'libertarian' attitudes to free speech. This abstentionist approach to moderation explains the lower volume of Christchurch content identified on these platforms because users are not required to repeatedly edit and re-upload their files to avoid content moderation.

FAR-RIGHT CONTENT PRODUCED BY DESIGNATED TERRORIST ORGANISATIONS

Key finding

Atomwaffen Division-branded content was by far the most prolific of any far-right terrorist group across the online spaces monitored by Tech Against Terrorism.

SUBMISSIONS AND ALERTS BY FAR-RIGHT TERRORIST GROUP

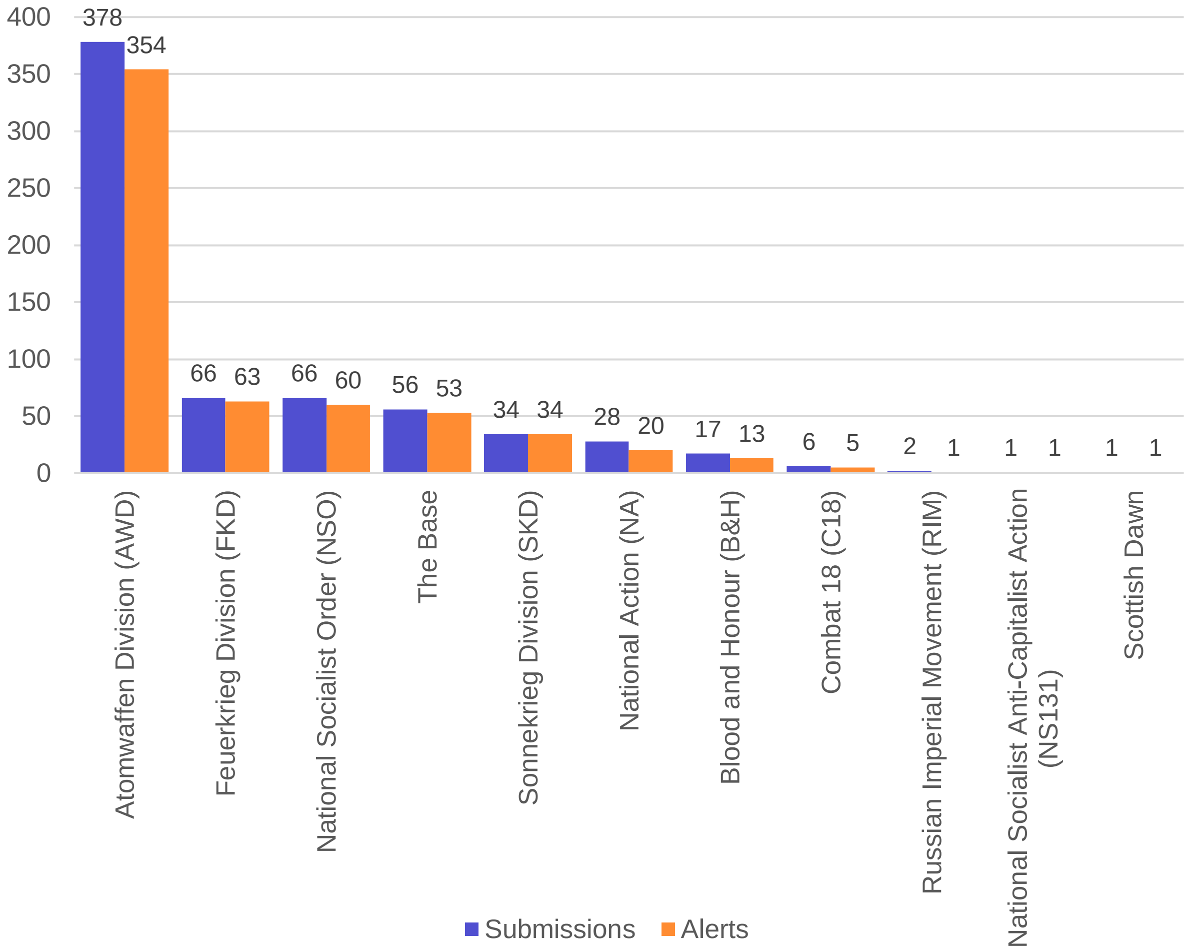


Figure 13: TCAP submissions and alerts by far-right terrorist group.

The volume of AWD propaganda content identified and submitted to the TCAP was far higher than the other designated far-right groups. This is likely to be due to the failure of the other groups, such as Feuerkrieg Division, to rival AWD as a credible neo-Nazi accelerationist brand. Many of these groups are either defunct, have split internally, or are not as active as when they were designated. For example, Feuerkrieg Division was active online in public channels until July 2022, since which time very little propaganda has been identified.

TCAP SUBMISSIONS BY FAR-RIGHT TERRORIST GROUP ACROSS PLATFORM TYPES

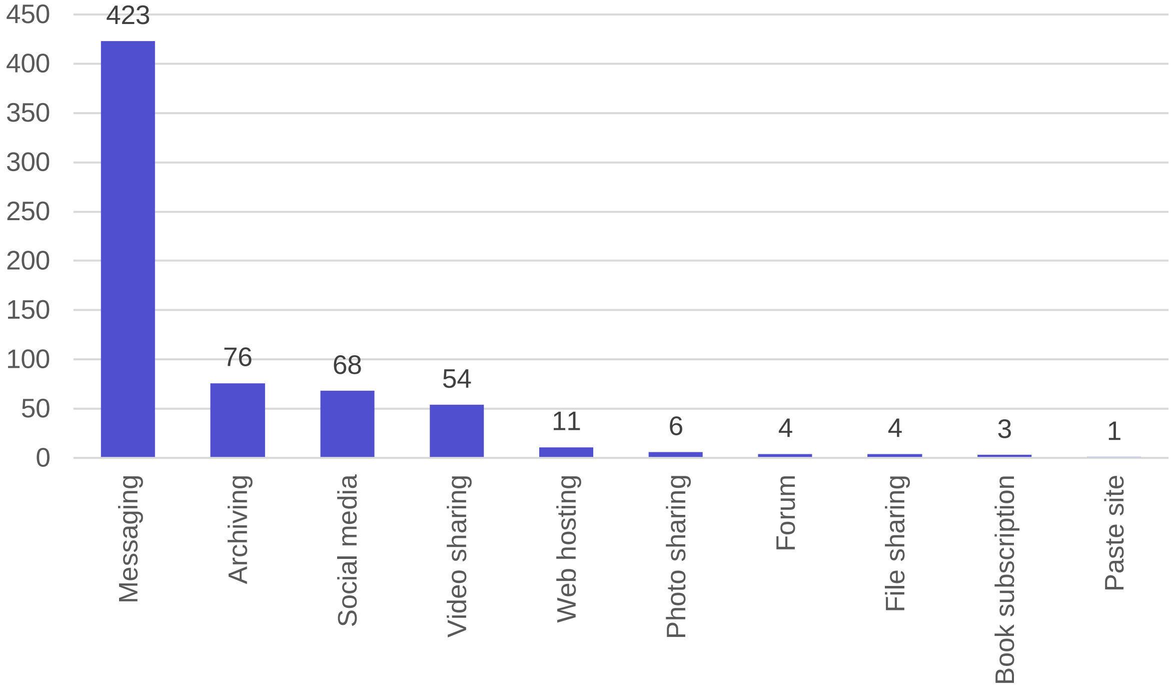


Figure 14: TCAP submissions by far-right terrorist group per platform type.

Far-right group content was heavily concentrated on one messaging platform which is popular with these groups because it offers both public and private channels, large file-size limits, and the ability to broadcast to sizeable audiences. Archiving sites provide a stable location for group propaganda, whereas alt-tech video-sharing sites tend to have more sympathetic audiences and lower levels of moderation of far-right content. Social media platforms continue to constitute an attractive target for supporters of far-right groups, given the larger potential reach for their propaganda.

PART 2: TECH PLATFORM RESPONSES TO FAR-RIGHT TERRORIST CONTENT

UNCOOPERATIVE OR HOSTILE PLATFORMS

Key finding

Smaller alt-tech video-sharing and social media platforms are being exploited by extremists to host far-right terrorist content. Tech Against Terrorism is unable to engage with these platforms to alert them to this content.

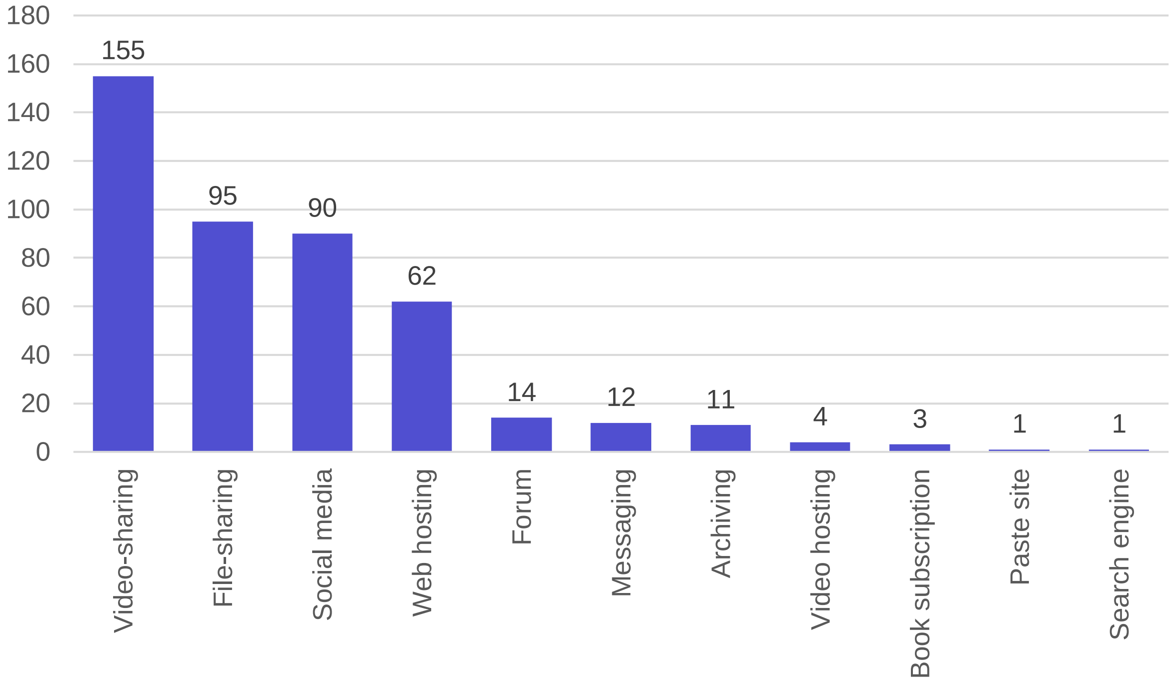


Figure 15: Far-right submissions across platforms not alerted by TCAP, by platform type.

Out of 2,966 URLs containing far-right terrorist content which were identified and submitted to the TCAP, 448 URLs were not alerted due to being hosted on platforms that do not receive TCAP alerts. This non-alerted content comprises 15% of the total far-right submissions and was spread across 84 different platforms, many of which were small file-sharing sites, forums, or websites. We do not currently issue alerts to these platforms for a multitude of reasons, but primarily because 'libertarian' attitudes to free speech, or in some cases an ideological affinity with the content, make them unwilling to engage on matters of content moderation. Other challenges to engagement include the nature of the platform (if it is a website[9] or forum) and the location of the platform (if it is hosted in an authoritarian state).

WHY ARE SOME PLATFORMS HARD TO ENGAGE?

Far-right networks have migrated towards more niche, alt-tech platforms in the past few years as larger platforms have become more stringent in their moderation of far-right content. Some of the 'free speech' video-sharing platforms have attracted large numbers of far-right extremists due to their resistance to strict content moderation and their offer of features similar to larger platforms such as YouTube. Alternative social media platforms have a similar 'free speech' outlook which encourages exploitation by far-right actors who have been able to post terrorist content on their services without removal. We have identified a significant volume of far-right terrorist content on these two platform types (245 URLs), none of which we have been able to alert for the reasons outlined above.

A smaller volume of far-right terrorist content has been identified across a wide number of file-sharing platforms, far-right and gore websites, and fringe forums. This does not accurately reflect the volume of content likely to be hosted on these sites; we do not prioritise identifying content on these more extreme platforms given the unlikelihood of this content being removed.

REMOVAL RATES OF ALERTED FAR-RIGHT CONTENT

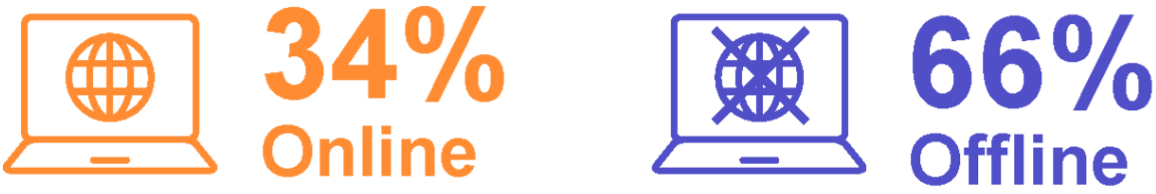


Figure 16: Overall tech platform removal rate of far-right terrorist content alerted through the TCAP.

As of 31 October 2023, around two thirds of far-right terrorist content alerted through the TCAP had been removed by tech companies and is now offline. This is low relative to the removal rate of Islamist content alerted through the TCAP, which is currently 78%. In year 1 of the TCAP, the removal rate of far-right content was 50% (out of 115 alerts), rising to 61% in year 2 (out of 738 alerts).[10] Despite the massive increase in far-right alerts sent since the end of year 2 (November 2022) with a total of 2345 alerts sent as of 31 October 2023, the removal rate has remained relatively stable at 66%.

[9] Our Open-source Intelligence (OSINT) team disrupts terrorist-operated websites by engaging with infrastructure providers outside of the alerting framework of the TCAP.

[10] Tech Against Terrorism, TCAP Transparency Report 2021 – 22. https://terrorismanalytics.org/policies/transparency-report

As outlined in a previous blog, there are multiple reasons for the relatively low removal rate of far-right content. The historical prioritisation of violent Islamist terrorism in counterterrorism legislation means the violent far-right has often been overlooked, which has resulted in confusion over the legality of this content for tech platforms. This in turn means that smaller tech companies in particular may lack the knowledge and capacity to understand, detect, and moderate far-right terrorist content. Furthermore, far-right terrorist content is inherently more difficult than Islamist content for tech companies to identify because doing so requires a high degree of expert knowledge of far-right symbols and phrases.

A significant proportion of the far-right content alerted through the TCAP is hosted on alt-tech video-sharing or social media platforms. These platforms have a higher threshold for content removal due to 'libertarian' policies on freedom of speech in line with US laws, which possibly explains lower takedown rates of alerted TCAP URLs. There are also jurisdictional gaps: for instance, content supporting a UK-proscribed group may be illegal in the UK but not in the US where the tech company is based and whose laws focus on incitement to violence. Therefore, tech companies may only ban content in particular jurisdictions, meaning it remains accessible elsewhere, or domestically with the use of a VPN.

HOW DO TECH PLATFORMS DIFFER IN THEIR MODERATION OF FAR-RIGHT CONTENT?

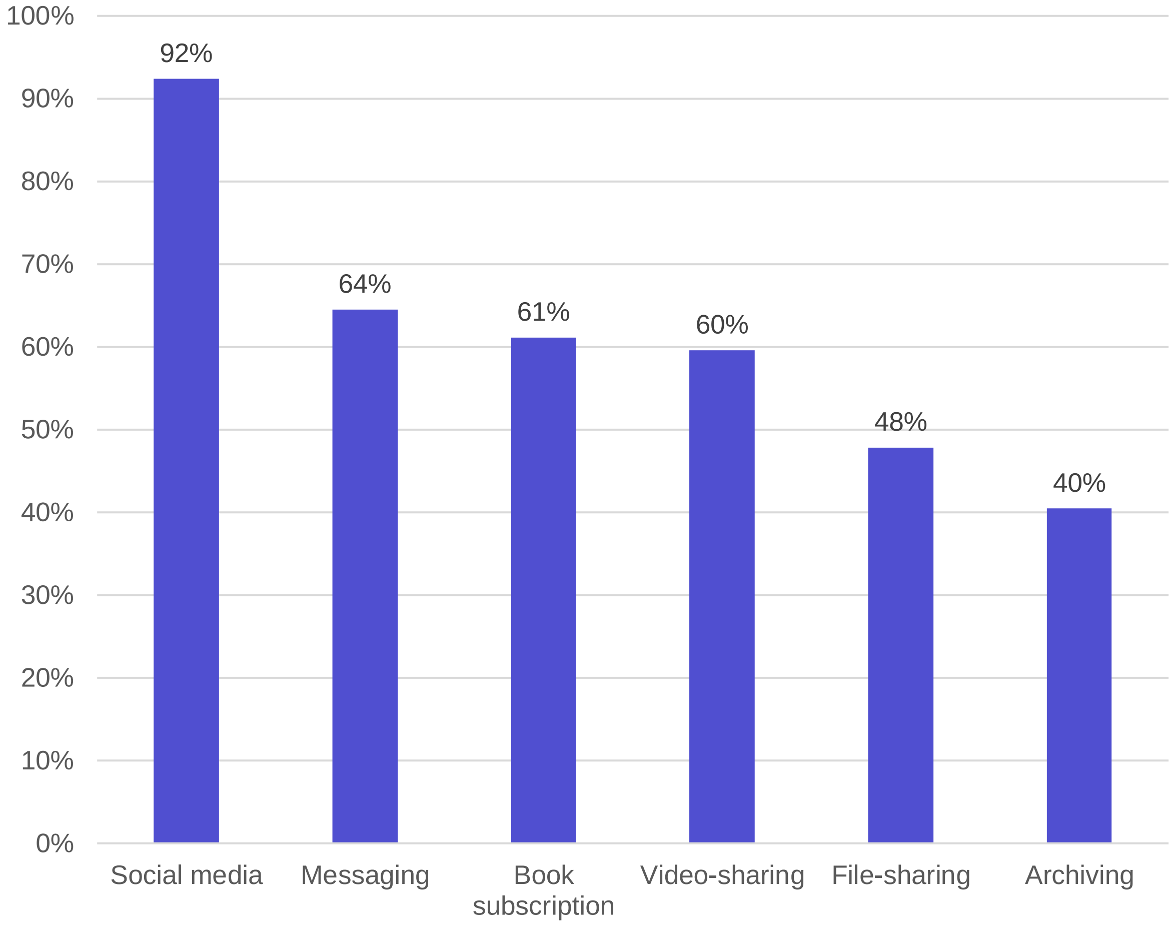


Figure 17: Average removal rate of alerted far-right terrorist content by tech platform type. NOTE: Only includes platform types which received over 10 TCAP alerts.

The removal rate of far-right terrorist content varies significantly by tech platform type. The data shows that social media platforms, generally those with a larger number of users and a higher profile, are more likely to remove alerted material. Reasons for this are likely to include: more comprehensive content moderation policies covering a wider range of perceived harms; greater compliance with jurisdictional legislation due to a global userbase; the reputational risk of hosting content advocating racist ideas or depicting acts of terrorism; and a greater capacity to moderate content effectively with larger moderation teams or more sophisticated automation.

It is more difficult to generalise as to why content is or is not removed from other platform types. The content moderation decisions taken by individual platforms respond to the specific context in which they operate; this may, for example, lead to two file-sharing platforms having widely different enforcement rates of far-right terrorist content. This factor is even more prominent given that the far-right terrorist content we identified was concentrated on relatively few platforms. With that caveat, we make the following assessments:

- **Messaging platforms**: Given the heavy concentration of far-right content identified on messaging apps (See Fig. 14), their response to TCAP alerts has a disproportionate impact. One particularly prominent messaging platform is targeted by far-right actors because it offers significant opportunities to network, allowing a range of users to discover and receive material, while also providing a more secure environment with less restrictive content moderation policies. This platform's terms of service are unclear in relation to what far-right content violates their policies, and the nature of their internal decision-making remains opaque.

- **Video-sharing platforms**: The average removal rate of video-sharing platforms was 60%, slightly lower than for messaging platforms. The community guidelines of some of these platforms position them as more committed than 'Big Tech' to the free exchange of ideas and tend to focus on prohibiting illegal content. However, we have identified inconsistent enforcement of these policies with identical content in some cases removed and in others remaining online for extensive periods. This inconsistency is likely to be caused by confusion about the illegality of specific far-right content, what is entailed by removal orders from internet referral units or law enforcement, and inefficient moderation systems both for reviewing flagged material and recording previous decisions.

- **Archiving platforms**: Archiving services are heavily exploited by terrorist and violent extremist (TVE) actors across the ideological spectrum. In a far-right context, this exploitation means that a huge array of files are directly uploaded to these sites, and especially documentary written records (such as attacker manifestos or the textual matter of James Mason). The Terms of Service of archiving platforms tend to be minimalistic and focus on the user's responsibility for the content they upload. The lack of detailed public policies on what content is allowed on these services has made it impossible to infer why some content subject to TCAP alerts is removed while other content is not. It is likely that the lack of far-right entities designated as terrorist by the U.S. government minimises the legal incentives to moderate this content.

HOW IS DIFFERENT FAR-RIGHT CONTENT MODERATED BY TECH PLATFORMS?

Figure 18 breaks down by entity the removal rate of far-right terrorist content alerted via TCAP to tech platforms. The variation between entities demonstrates the fact that legal tools such as designation or classification as 'Objectionable' do not provide sufficient rationale for platforms to remove certain kinds of content (assuming platforms are aware of the legal status of entities in the first place).

REMOVAL RATE BY FAR-RIGHT TERRORIST ENTITY

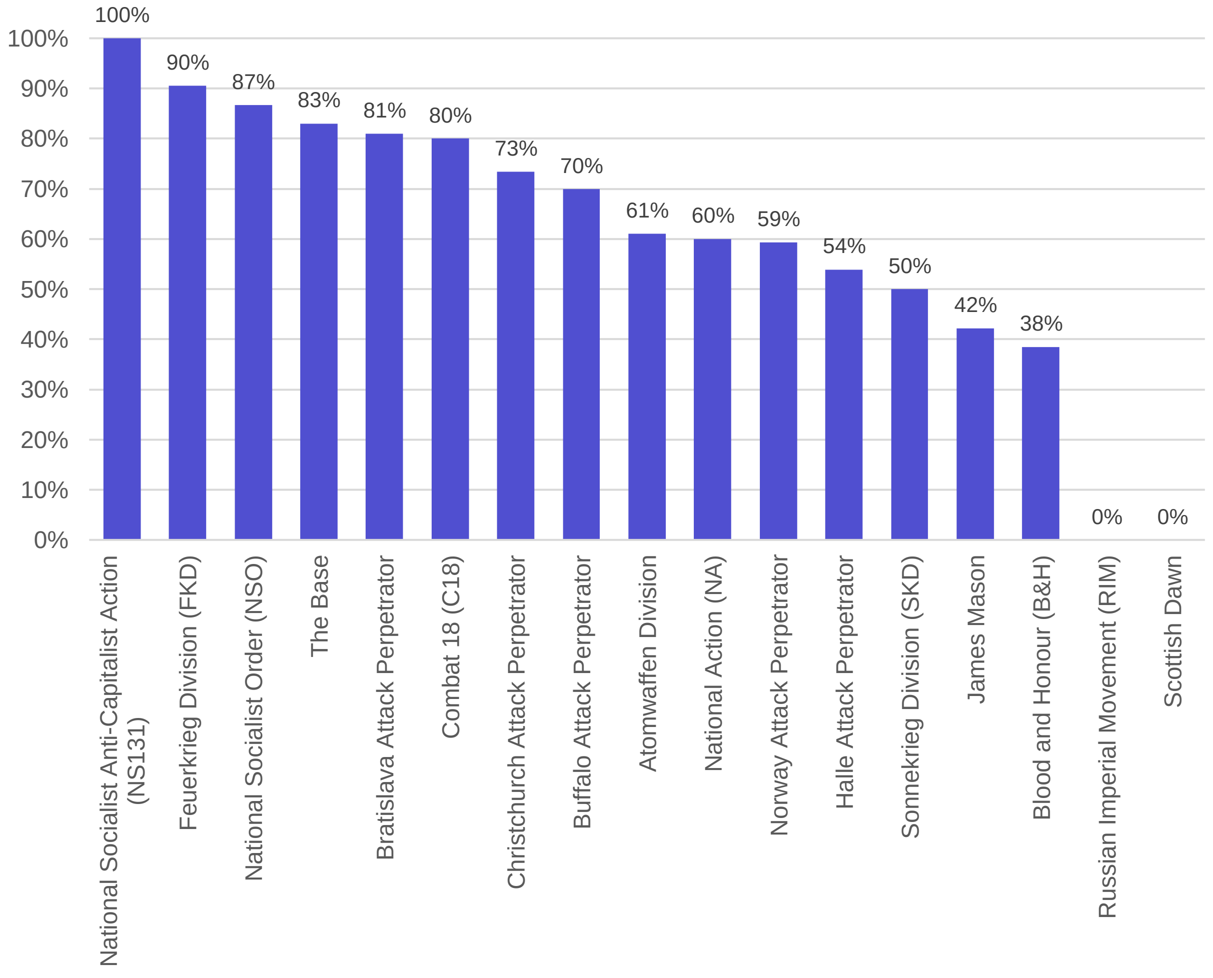


Figure 18: Average removal rate of alerted far-right content by terrorist entity.

Far-right terrorist groups

There is plainly a wide divergence between the removal rates of content produced by different far-right terrorist organisations. For certain groups, such as Russian Imperial Movement or NS131, small sample sizes of content make meaningful assessments impossible. For the other organisations, it remains difficult to establish the reasons for variations in the enforcement of content moderation, although contributing factors are likely to include where the content is hosted (and the platform's policy on moderating content based on location) and the nature of the content itself (such as how graphic it is) in addition to any specific group with which the content may be affiliated.

Far-right terrorist attackers

Of the individual far-right attackers whose content is alerted by TCAP, none is removed more than 81% of the time (the Bratislava attacker) or less than 54% of the time (the Halle attacker). The relatively high removal rate of the Bratislava manifesto may be due to the reduced spread of this content as well as the recency of the incident exposing the content to improved moderation processes.

The attacks in Christchurch, New Zealand, and Buffalo, New York, received huge and sustained levels of media attention. The presence in both attacks of a protracted and extremely violent livestream - in addition to a written manifesto – makes removal of content associated with the attacker more likely. It is highly likely that the distressing nature of the videos associated to these attacks which display murder justifies widespread removal across platforms.

The Halle attacker also produced a livestream and manifesto. It is realistically possible that the relatively small number of victims has contributed to the low salience of this attack amongst the wider public, including platform moderators and hence a lower removal rate.

Finally, the Norway attacker also engaged in a mass shooting but produced no video record. The perpetrator’s manifesto, a document more than 1,500 pages in length, was produced almost a decade before the Christchurch attack. The novel nature of such a document in 2011, as well as its extreme length and the extensive ‘philosophy’ behind the crime, has meant that, for years after the attack, legitimate news outlets shared it prolifically. It is a realistic possibility that the legacy of this attention has made it more likely for content moderators to perceive the document to have a residual historical or political interest.

James Mason

As the only far-right individual in scope who has *not* conducted a terrorist attack, James Mason occupies a unique space in TCAP alerts. Mason’s connection to far-right terrorism is well established, and he has produced a vast corpus of content including books, pamphlets, essays, and videos.[11] Despite this, it is likely many platforms are unaware of Mason’s links to terrorism and violent Neo-Nazism.

Given Mason is not designated as a terrorist by the United States, his home country, and there are no legal restrictions on his ability to disseminate his views, it is possible that some US-based platforms consider his self-produced content to be within the limits of permissible political speech. This is especially likely to be the case for tech platforms with more ‘libertarian’ outlooks. Furthermore, since Mason’s content is not visually violent or gory, platforms may overlook its dangerous links to violent extremism.

[11] https://www.terrorismanalytics.org/project-news/new-TCAP-entity-James-Mason

2. SPOTLIGHT: TRENDS IN FAR-RIGHT TERRORIST EXPLOITATION

KEY TRENDS

FAR-RIGHT TERRORIST LIVESTREAMS

In recent years, far-right terrorists have increasingly exploited online platforms to livestream their acts of violence and/or disseminate attacker-produced content that aims to justify or incite others to commit violence. A terrorist attack on two mosques in Christchurch, New Zealand, in which 51 people were killed, represented the first notable instance of a far-right inspired terrorist attack being livestreamed online. The perpetrator livestreamed the attack on Facebook, with the video being subsequently reposted across the internet reaching a massive audience.

More recently, an attacker killed 10 and injured 3 others in a shooting at a supermarket in Buffalo, New York, USA. The perpetrator livestreamed the attack on Twitch, released a manifesto detailing their motivation for the attack on Google Drive, and posted a transcript of an online diary originally hosted on Discord. Copies of the livestream, manifesto, and online diary which were produced by the perpetrator spread rapidly across the internet.

FAR-RIGHT TERRORIST MANIFESTOS

Tech Against Terrorism has also observed the online spread of manifestos produced by far-right perpetrators of real-world acts of violence, both recent and historic. These manifestos are typically used to justify the perpetrator’s violent actions, propagate their worldview, and provide an instructional template for further violence. Since its inception, the Terrorist Content Analytics Platform (TCAP) has expanded to now alert the manifestos from 6 different terrorist attack perpetrators.

THE TERRORGRAM NETWORK

“Terrorgram”[12] is a collective formed on the messaging app Telegram that creates and disseminates propaganda supporting terrorism and militant accelerationism.[13] This loosely affiliated neo-fascist network of channels has produced branded propaganda since at least 2019 and has been closely linked to designated far-right entities including Atomwaffen Division (AWD), The Base, Feuerkrieg Division (FKD), and James Mason. The messaging on Terrorgram has influenced far-right extremists, including the perpetrator of a mass shooting at an LGBTQIA+ bar in Bratislava, Slovakia in 2022 who specifically acknowledged the influence of Terrorgram in his manifesto.[14]

SANCTIFICATION

This refers to a cultural trend in fringe violent far-right online communities which venerates the violent actions of far-right extremists.[15] The contemporary referencing of terrorists as Saints

[12] “Terrorgram” comprises a network of far-right TVE actors operating tens of messaging channels, primarily on Telegram. The network has been producing propaganda since at least 2019, promoting a narrative that is overtly supportive of terrorism, and other forms of political violence, to further their militant accelerationist goals. Associated channels have been frequently suspended since 2022.

[13] Source: https://gnet-research.org/2022/09/12/analysing-terrorgram-publications-a-new-digital-zine/

[14] Source: https://www.rsis.edu.sg/rsis-publication/icpvtr/bratislava-shooting-the-making-of-terrorgrams-first-saint/

[15] Source: https://gnet-research.org/2023/04/27/the-lineage-of-violence-saints-culture-and-militant-accelerationist-terrorism/

began with the perpetrator of the Christchurch attack who was quickly anointed “Saint Tarrant”, with a wave of far-right attackers thereafter ‘sanctified’ to form a broader shrine to accelerationist violence. The Terrorgram community has been one of the key drivers of “Saints Culture” expanding the “calendar”[16] of Saints to over 50 individuals and creating propaganda that glorifies them and incites accelerationist violence.

GAMIFICATION

Another notable trend of emerging exploitation includes the gamification of real-world acts of violence inspired by far-right extremist ideologies and grievances. Gamification includes the recreation of violence based on attacker-produced livestreams, including user-generated content to create virtual depictions of the attacks on gaming platforms aimed at children.

EXPLOITATION OF ALT-TECH PLATFORMS

A notable trend in the dissemination of far-right terrorist propaganda is the exploitation of ‘alt-tech’ platforms where counterterrorism policies and enforcement are either permissive or rudimentary. There are several alt-tech video-sharing sites which are routinely targeted by these actors likely due to their policies of ‘free speech.’ We have also observed far-right terrorist entities and networks operating accounts across several platforms simultaneously. This is likely a strategy to mitigate the impact of account removal and to maximise the number of online users viewing their content.

OFFLINE EVENTS

RUSSIAN INVASION OF UKRAINE

Violent far-right propaganda is closely shaped by the ongoing and unstable offline threat environment. Notably, the Russian invasion of Ukraine in 2021 drove significant debate among far-right extremists online. Russia justified the invasion as the “denazification” of Ukraine, which galvanised some far-right violent extremist users and confused others. A significant proportion of pro-Kremlin channels and networks frequently espoused long-standing extremist far-right rhetoric, such as antisemitism. Alongside anti-western and anti-liberal sentiment, this drove support for Russia from elements of the far-right. However, the perception of Russia as neo-Bolshevik and Ukrainians as defenders of ‘white Europe’ has won Ukraine sympathy among other far-right extremist online users.[17]

HAMAS-ISRAEL CONFLICT

More recently, Tech Against Terrorism has been monitoring the online threat landscape in response to the Hamas-Israel conflict. This includes observing and analysing far-right extremist discourse surrounding the incident. The primary narratives that we have observed have focused on how a conflict between Jewish and Muslim communities is likely to impact White people in the future. Tech Against Terrorism has identified far-right narratives including conspiracy theories around the Hamas terrorist attacks being an Israeli orchestrated ‘false flag’, the celebration of violence against Jewish and Muslim communities, and heightened antisemitism and Islamophobia in western countries and the incitement of violence against pro-Palestine protesters.

[16] Telegram channels dedicated to Saints published monthly calendars celebrating the dates of far-right attacks and marked other milestones including the arrests, deaths and even birthdays of perpetrators. Propagandists also produced fact sheet graphics that broke down the details of each attack into easily readable paragraphs alongside the death toll and the number of wounded, attack method and status of each killer alongside a montage of their photos, pictures of their weapons victims and crime scenes. Source: https://gnet-research.org/2023/04/27/the-lineage-of-violence-saints-culture-and-militant-accelerationist-terrorism/

[17] More information can be found in Tech Against Terrorism’s 2022 State of Play report.

SPOTLIGHT

In this section, we develop our analysis of several of the key trends identified above by providing examples of how these threats manifest in practice. The notable trends and developments we discuss below were identified through the daily monitoring of online spaces by Tech Against Terrorism's open-source intelligence team throughout 2023.

GAMIFICATION

Since expanding the TCAP's remit to include 'promotional' terrorist material in July 2023,[18] Tech Against Terrorism has identified 141 URLs containing 'gamified' versions of far-right terrorist attacks and sent 117 alerts to 5 different companies.[19] This is content that utilises the world-building potential of games to recreate the actions and scenes from a real attack livestream. On at least one occasion, this was synthesised with part of the actual livestream also playing in the same frame. Other instances involve the use of modifications or 'mods' from the Steam app to recreate the setting of a mass shooting. The technical sophistication and attention to detail of these recreations vary greatly.

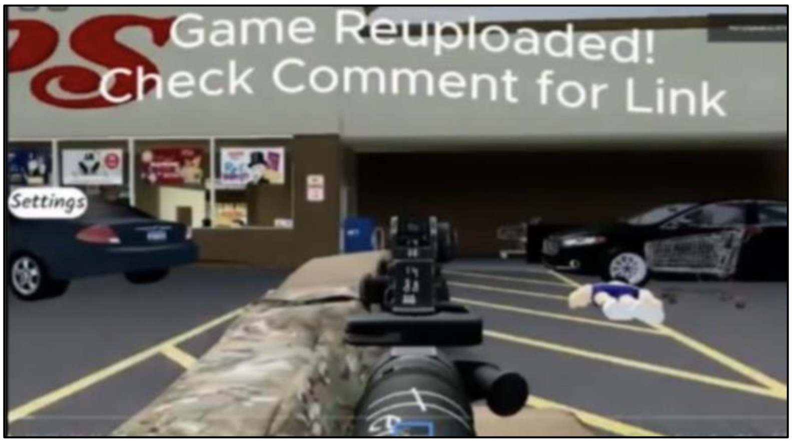


Figure 19: Gamified edit of the Buffalo attack livestream. Screenshot captured 17 January 2024.

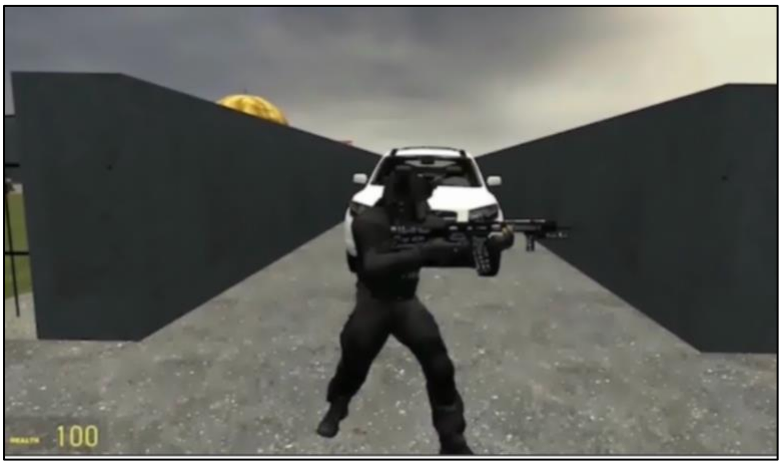
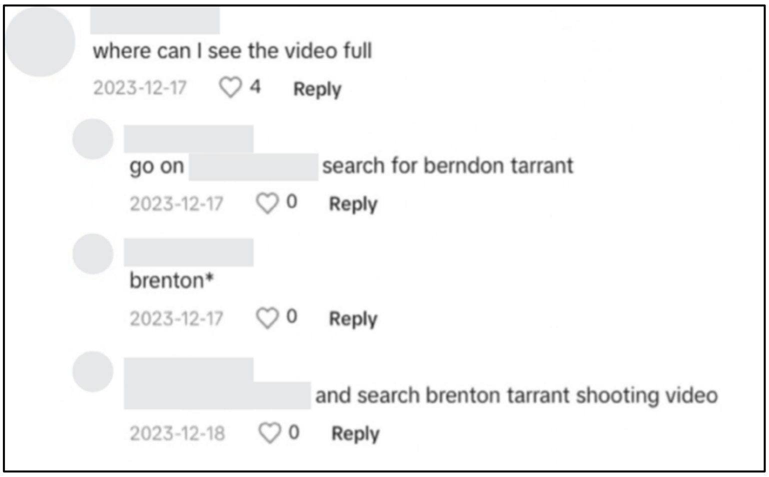


Figure 20: Gamified character depicting the Christchurch attacker. Screenshot captured 7 February 2024.

[18] Source: https://www.terrorismanalytics.org/project-news/TCAP-Tiered-Alert-Launch

[19] This statistic was recorded on 1 March 2024.

Tech Against Terrorism regularly observes such content being disseminated across both mainstream social media platforms and secure messaging spaces. The prevalence of this content and its creator networks on mainstream social media platforms is especially concerning given the potential for wider audience reach as well as the difficulty these platforms have experienced in identifying and removing this content. We are also concerned that the sanitised and accessible nature of these violent games can act as a gateway to more extreme content and views, with younger users especially vulnerable. Our analysts regularly identify examples of users seeking and being signposted to more extreme content (normally unedited livestreams) elsewhere on the internet within the comments sections of these videos (See *Fig. 22*).



[redacted]
where can I see the video full
2023-12-17 ♡ 4 Reply

[redacted]
go on [redacted] search for berndon tarrant
2023-12-17 ♡ 0 Reply

[redacted]
brenton*
2023-12-17 ♡ 0 Reply

[redacted] and search brenton tarrant shooting video
2023-12-18 ♡ 0 Reply

Figure 21: User requesting information linking to a video source in the comments section of a gamified edit. Screenshot captured 21 January 2024.

It is highly likely that gamified content will continue to proliferate and that the ‘meme culture’ inherent to far-right online activity will ensure this content is shared as widely as possible on mainstream platforms. Tech Against Terrorism has also observed an increased trend of users commenting on one another’s ‘edits’ of original video material with approval or compliments – some users have even watermarked their content with a handle as a form of signature. Such creators can develop a particular ‘style’ of imagery, with curated visual effects repeatedly applied or exhibiting an apparent specialism in individual attack livestreams.

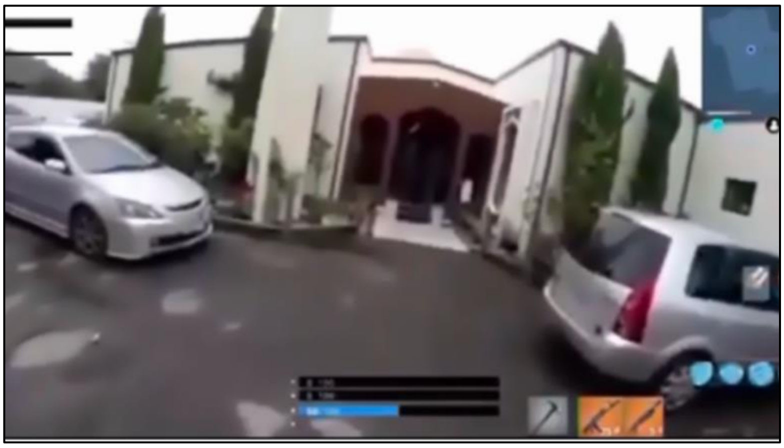
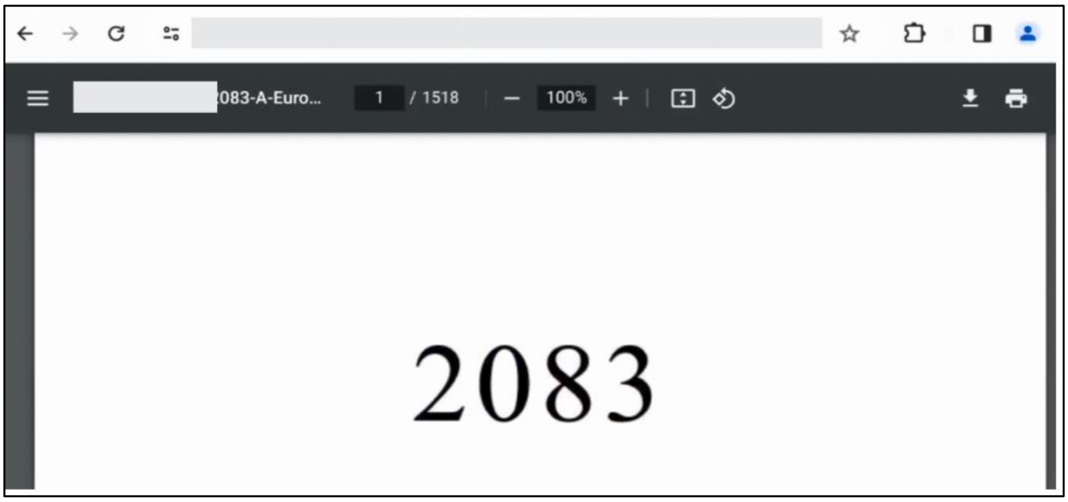


Figure 22: Footage from Christchurch attack livestream edited and overlaid with graphics intended to resemble the video game Fortnite. Screenshot captured 8 February 2024.

HOSTING OF MANIFESTOS

Although the majority of far-right TCAP alerts notify social media platforms and secure messaging apps, we have also observed the hosting of far-right terrorist material in several unusual or unexpected internet locations.

On 7 occasions throughout July 2023, Tech Against Terrorism notified media organisations that a copy of the Norway attacker's manifesto was hosted on their site. As of February 2024, only 2 of the relevant organisations had removed the content. These media organisations represented a range of countries across Europe and included nationally known broadcasters or print titles. As discussed previously, it is likely that the then-novel nature of the Norway perpetrator's manifesto in 2011 led media organisations to host versions of the document for its 'news value'. Such copies continue to surface over a decade later and are indexed via search engine results.

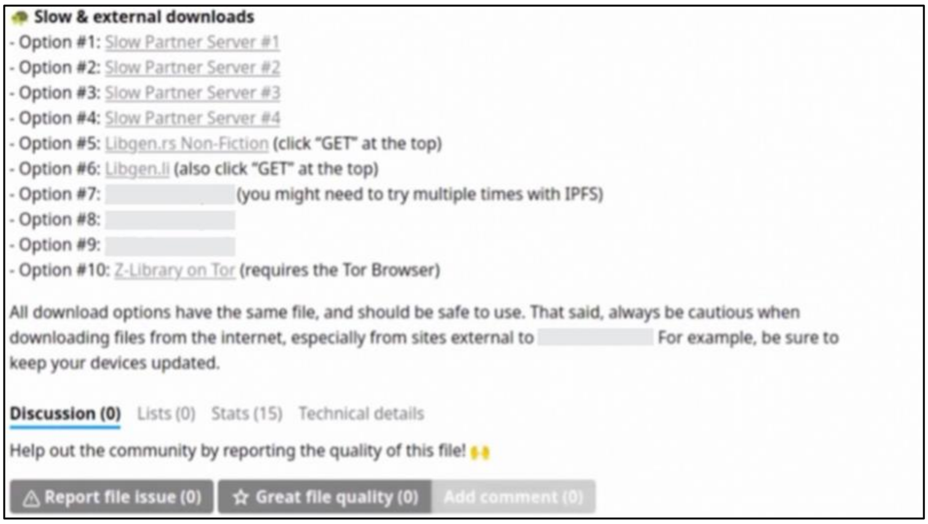


2083

Figure 23: Redacted screenshot of Norway attacker manifesto hosted in full on the website of a major national broadcast news channel. Screenshot captured 8 February 2024.

In 2023, we identified several copies of far-right terrorist manifestos hosted in online code repositories and on translation sites. The contribution of multilingual users to the proliferation of far-right terrorist content is highlighted by three consecutive posts on a Russian social media site which included 11 versions of the Christchurch attacker manifesto, from Bulgarian to Dutch.

Archiving sites have played a prominent role in the hosting of far-right material as the content can be shared via a simple URL. Through TCAP we have identified far-right content on 8 different sites that could be described as 'archiving' or 'library' in nature. One archiving site hosted dozens of copies of multiple far-right manifestos via gateways to other library sites and to the Inter-Planetary File-Sharing (IPFS) network.



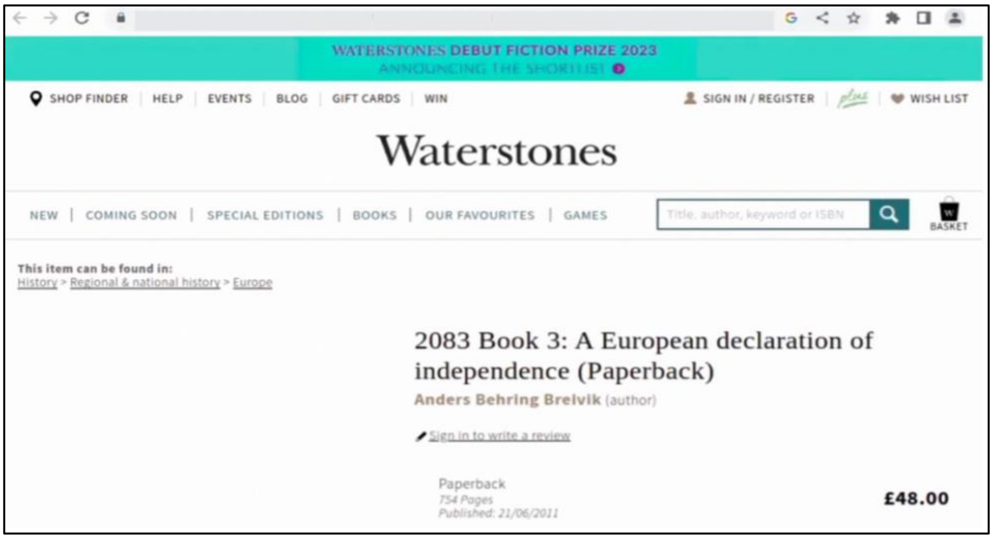
Slow & external downloads
- Option #1: Slow Partner Server #1
- Option #2: Slow Partner Server #2
- Option #3: Slow Partner Server #3
- Option #4: Slow Partner Server #4
- Option #5: Libgen.rs Non-Fiction (click "GET" at the top)
- Option #6: Libgen.li (also click "GET" at the top)
- Option #7: [redacted] (you might need to try multiple times with IPFS)
- Option #8: [redacted]
- Option #9: [redacted]
- Option #10: Z-Library on Tor (requires the Tor Browser)

All download options have the same file, and should be safe to use. That said, always be cautious when downloading files from the internet, especially from sites external to [redacted] For example, be sure to keep your devices updated.

Help out the community by reporting the quality of this file! 🙌

Figure 24: Screenshot shows aggregated links to Christchurch attacker manifesto on archiving site. Screenshot captured 6 July 2023.

Additionally, we found copies of manifestos hosted on websites dedicated to Christian fundamentalism, 'gun rights', and a website purportedly devoted to 'crime research.' We also identified copies of the Norway attack manifesto available to purchase on the websites of three major UK booksellers - an investigation which was covered in The Observer.[20] It is likely that the automated ingest of catalogued content led to this scenario rather than an active choice made to feature the document across three 'books' for purchase.[21] This further highlights the ability of certain kinds of far-right terrorist material to proliferate online due to its 'news value,' the variety of means in which it is presented, and a lack of awareness by non-user-to-user services of how their platforms can be exploited by hostile actors.



This item can be found in:
History > Regional & national history > Europe

2083 Book 3: A European declaration of independence (Paperback)
Anders Behring Breivik (author)

Sign in to write a review

Paperback
754 Pages
Published: 21/06/2011

£48.00

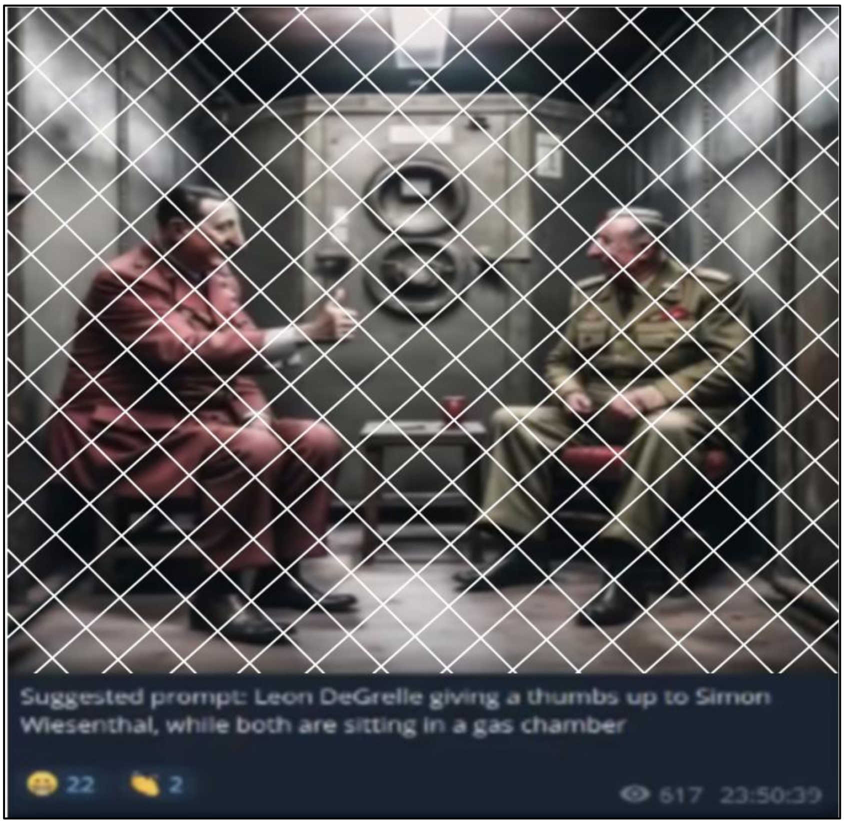
Figure 25: Waterstones listing for section of Norway attacker manifesto. Screenshot captured 19 July 2023. The listing has now been removed.

[20] Source: https://www.theguardian.com/world/2023/jul/23/terrorist-anders-breiviks-manifesto-was-listed-for-sale-on-waterstones-website

[21] Source: https://techagainstterrorism.org/in-the-news/terrorist-anders-breiviks-manifesto-was-listed-for-sale-on-waterstones-website

EXPLOITATION OF GENERATIVE AI

As we outlined in our report on early exploitation of Generative AI by terrorists and violent extremists, Tech Against Terrorism has noted that far-right actors have adopted AI tools to generate new visual messaging.[22] One channel we identified is dedicated to sharing neo-Nazi, antisemitic, and racist images generated by user 'prompts.' One such image purported to depict Leon Degrelle, a Waffen SS officer, and the 'Nazi-hunter' Simon Weisenthal sitting together in a gas chamber.



Suggested prompt: Leon DeGrelle giving a thumbs up to Simon Wiesenthal, while both are sitting in a gas chamber

Figure 26: Image apparently generated by an AI tool in response to the provided user prompt.

Such tools provide myriad opportunities for exploitation in the so-called 'shitpost' style of much far-right messaging, which is deliberately intended to shock while offering an occasional veneer of surrealism or mockery. The exploitation of AI art apps to generate far-right propaganda significantly reduces the effort required of users to produce creative and subtle memes by requiring only simple prompts. We have also identified guides that instruct creators on evading blocks applied to specific prompts and on generating imagery that evades detection by content moderation.[23]

[22] Source: https://techagainstterrorism.org/hubfs/Tech%20Against%20Terrorism%20Briefing%20-%20Early%20terrorist%20experimentation%20with%20generative%20artificial%20intelligence%20services.pdf

[23] Source: https://techagainstterrorism.org/hubfs/Tech%20Against%20Terrorism%20Briefing%20-%20Early%20terrorist%20experimentation%20with%20generative%20artificial%20intelligence%20services.pdf

RISKS OF FUTURE FAR-RIGHT EXPLOITATION OF GENERATIVE AI FOR VISUAL PROPAGANDA

Tech Against Terrorism has identified that the early experimentation with Generative AI indicates an emerging threat of TVE exploitation in the medium to long term, for the purposes of producing, adapting, and disseminating propaganda. Early research into the topic has identified several concerning use cases for Generative AI in relation to far-right terrorist propaganda.

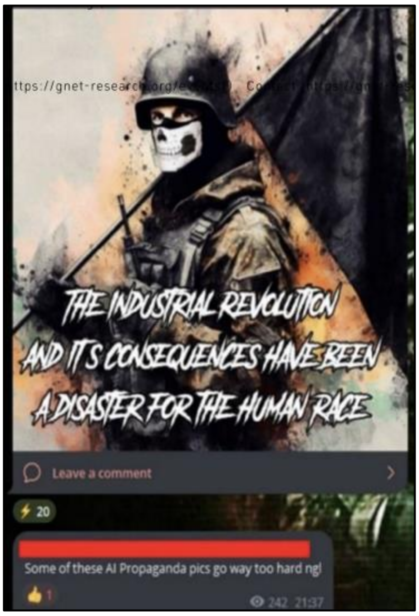
SANCTIFICATION

Generative AI is likely to augment existing efforts to glorify terrorist actors, whether living or dead. As sanctification has become important to fostering in-group identities within violent extremist networks online, it is highly likely to remain an integral element of future visual propaganda efforts. Such attempts have already been identified on one mainstream platform, in which the Christchurch perpetrator is depicted from a positive, pro-Ukrainian stance, with his idol-like status reinforced with the hashtag #brentontarranthero.

Figures 27 & 28: Images likely created using Generative AI of the Christchurch perpetrator. Screenshots captured 13 February 2024.

AI was used to propagate quotes by US domestic terrorist Ted Kaczynski, another revered Terrogram 'saint', on an accelerationist Telegram channel. Figure 29 displays the quotation "*the Industrial Revolution and its consequences have been a disaster for the human race",* taken from the so-called Unabomber's manifesto, in a distinctive Terrorgram and accelerationist aesthetic.[24] Thus, as the far-right experiment with AI, there is a realistic possibility that future terrorist propaganda uses synthetic text to quote 'saints' directly and thereby celebrate perpetrators and generate support.

[24] Source: https://gnet-research.org/2024/02/23/weapons-of-mass-hate-dissemination-the-use-of-artificial-intelligence-by-right-wing-extremists/


ttps://gnet-research.org/
THE INDUSTRIAL REVOLUTION
AND IT'S CONSEQUENCES HAVE BEEN
A DISASTER FOR THE HUMAN RACE

Leave a comment

20

[redacted]

Some of these AI Propaganda pics go way too hard ngl

1

242 21:37

Figure 29: 'A man in tactical fear, wearing a skull mask. The beginning of Ted Kaczynski's Manifesto was written with a font that can be easily associated with the Terrorwave aesthetic'. Source: [Global Network for Extremism and Technology](), 2024.

CIRCUMVENTION METHODS

EDITING

Generative AI supports creative editing processes that can subtly reinforce extremist ideals. On 4chan, AI generated images were embedded with right-wing extremist tropes and symbols, including the Happy Merchant, the Sonnenrad, and Adolf Hitler.[25] It is notable that these edits are designed to reinforce antisemitic beliefs subconsciously as well as to improve content evasion methods through obfuscation. Future far-right propagandists are likely to exploit the benefits of creative editing to augment radicalisation efforts through memetic warfare.

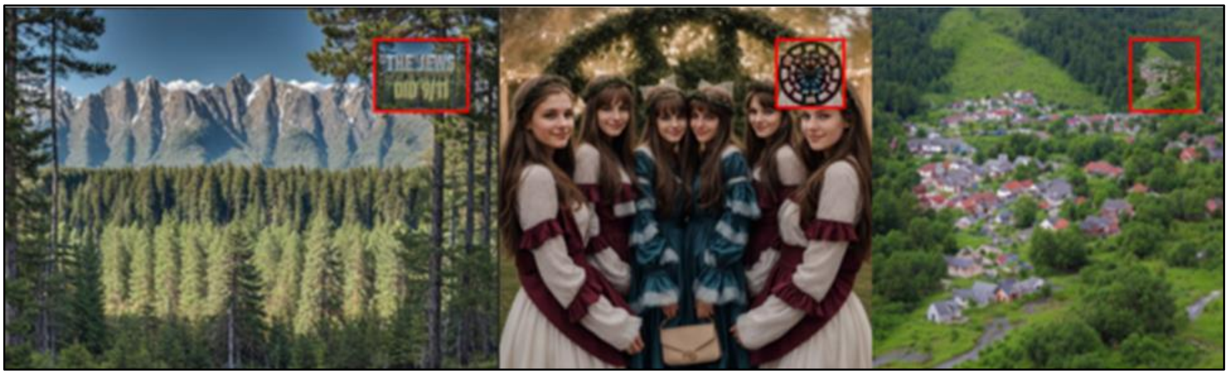


Figure 30: 'Antisemitic and Nazi imagery hidden in AI-generated images.'
Source: [Global Network on Extremism and Technology](), 2023.

[25] Source: https://gnet-research.org/2023/11/13/for-the-lulz-ai-generated-subliminal-hate-is-a-new-challenge-in-the-fight-against-online-harm/

MEDIA SPAWNING

Media spawning, in which multiple edits of the same content are created deliberately, can limit the efficacy of hash-matching techniques and thus slow the identification of content. The limitations of hash-matching are already apparent in the identification of terrorist propaganda that is not generated by AI[26] and constitute an existing challenge for the tech sector. Generative AI is highly likely to exacerbate this challenge with its scope for quickly and easily mass-producing image variations, thus burdening hash-detection systems further.

VARIANT RECYCLING

Finally, evidence suggests that existing propaganda can be recycled, either in its entirety, or using specific features such as logos, symbols, and images. There is a realistic possibility that generative AI helps terrorists to repurpose propaganda to create newer versions and subvert hash-detection mechanisms simultaneously. Telegram's far-right communities have already employed AI to resurface well-known existing symbols of extremism, including Pepe the Frog and Hitler.[27] This suggests that AI is likely to prove useful to terrorists should they desire to revamp existing propaganda by recycling key iconographic features. In this way, generative AI reduces the need for novel material in creative innovation and thereby supplements wider evasion tactics.

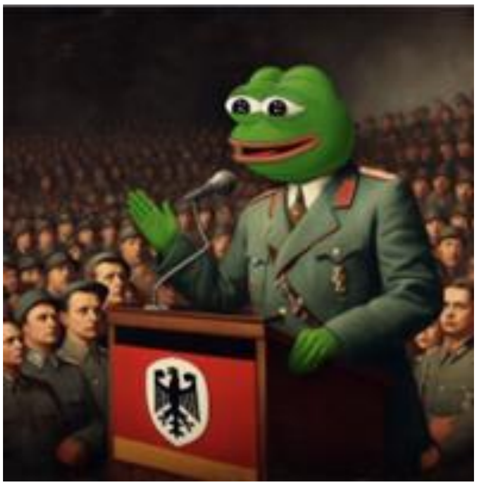


Figure 31: 'Pepe resembling Adolf Hitler.' Source: Global Network on Extremism and Technology, 2023.

EVENT-SPECIFIC PROPAGANDA AND DEEPFAKES

The conflict in Gaza has demonstrated that geopolitical events can trigger the creation of context-specific visual propaganda. Notably, research has identified generative AI efforts to depict the Israel-Hamas conflict in an antisemitic light on 4chan.[28] Figure 32 depicts a Jewish Star of David ablaze above the ruins of Gaza and is used to suggest Jewish support for the ongoing bombings and atrocities in the region. Generative AI is thus likely to enable terrorists to produce propaganda as conflicts arise, specifically exploiting these opportunities to broadcast their narratives rapidly, and at scale.

[26] Source: https://www.ofcom.org.uk/__data/assets/pdf_file/0036/247977/Perceptual-hashing-technology.pdf

[27] Source: https://gnet-research.org/2024/02/23/weapons-of-mass-hate-dissemination-the-use-of-artificial-intelligence-by-right-wing-extremists/

[28] Source: https://gnet-research.org/2023/11/13/for-the-lulz-ai-generated-subliminal-hate-is-a-new-challenge-in-the-fight-against-online-harm/

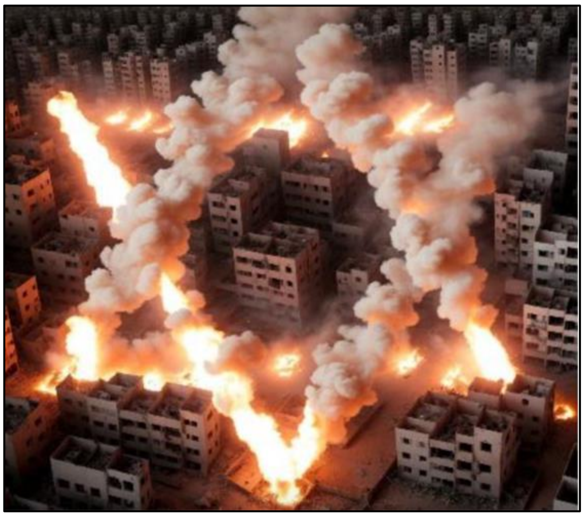


Figure 32: 'AI generated Jewish Star of David of bombing in Gaza.'
Source: Global Network on Extremism and Technology, 2023.

In addition, deepfake technology is highly likely to bolster future misinformation and disinformation efforts by terrorist actors. Research anticipates that far-right extremists will leverage deepfakes to undermine trust in government, as well as ideological 'out-groups'.[29] This could manifest in diverse forms, from creating 'proof' to validate causes through to calls for violence to remove governments using fabricated information.

[29] Source: https://www.icct.nl/sites/default/files/2023-12/The%20Weaponisation%20of%20Deepfakes.pdf

3. POLICY RECOMMENDATIONS FOR TECH PLATFORMS

UNDERSTANDING THE THREAT

⇒ We recommend that tech companies monitor for far-right terrorist-affiliated symbols and key words to improve detection of far-right terrorist content. Tech Against Terrorism's Knowledge Sharing Platform (KSP) provides an easily navigable and extensive database of far-right symbols and phrases for platform moderators. Access it here.

LEGAL CLARITY

⇒ We recommend that tech companies prohibit both terrorism and violent extremism in their Terms of Service / Community Guidelines.[30] A prohibition of terrorism should include content which encourages, supports, glorifies and/or promotes terrorism, terrorist organisations, terrorist attacks and attackers.

⇒ We recommend consulting national and supranational designation lists as a guide to the far-right entities that have been designated as terrorist through a stringent legal process.[31] Find our list of designated far-right groups here (and in Fig. 2).

⇒ We recommend that tech companies focus on content produced by far-right terrorist attack perpetrators, such as the manifesto and livestream produced by the Christchurch attacker. Find attacker-produced content included in the TCAP here.

TRANSPARENCY

⇒ We recommend that tech companies explain what is prohibited on their services in a way that is clear and easily understandable for users.

⇒ We recommend informing users why action has been taken against their content or account, and on what grounds, with reference to a specific policy violation.[32]

⇒ We recommend that tech platforms produce a transparency report about their moderation enforcement actions.

[30] The prohibition of violent extremism allows platforms to more effectively moderate violent far-right content that has not been produced by designated far-right groups.

[31] Canada, the UK, and Australia have the most developed designation lists for far-right entities. For more information on international designations systems and the implications for online terrorist content, see Tech Against Terrorism's report, 'Who Designated Terrorism? The Need for Legal Clarity to Moderate Terrorist Content Online'

[32] Tech platforms are encouraged to provide an explanation of why a user's content has been removed under the EU's Terrorist Content Online (TCO) Regulation and the Digital Services Act (DSA).

IMPROVING ENFORCEMENT OF TCAP ALERTS

A key finding from our analysis of tech platform responses to far-right terrorist content was that the overall removal rate of this content, relative to Islamist terrorist content, is low. The reasons for this vary across tech platform types, sizes, and content moderation approaches. However, we offer below **general recommendations** for tech platforms that respond to the specific challenges posed by far-right TVE content.

UNDERSTANDING THE THREAT

Our data shows that the removal rates of far-right TCAP alerts vary widely depending on the terrorist entity that produced the content.

It is likely that this stems from the challenge moderators face in identifying and contextualising the terrorist nature of the content. This could explain why the removal rate of far-right content relating to less well-known entities such as James Mason (a far-right ideologue) or Scottish Dawn (a little-known offshoot of National Action) is relatively low.

Another related factor is the challenge faced by smaller tech companies in particular in identifying far-right content because it is often not clearly affiliated to designated groups through branding (unlike Islamist terrorist content) and requires expert knowledge of far-right symbols and phrases.

Recommendations:

- Tech Against Terrorism recommends tech companies to monitor for far-right terrorist-affiliated symbols and key words to improve detection of far-right terrorist content. Tech Against Terrorism's Knowledge Sharing Platform (KSP) provides an easily navigable and extensive database of far-right symbols and phrases for platform moderators. It is accessible here. Further resources include the Anti-Defamation League's (ADL) Hate Symbols Database, Glossary of Extremism and Hate, as well as the Southern Poverty Law Center's (SPLC) extensive resources.

- The KSP also hosts a database of information on the far-right terrorist entities that we alert through the TCAP. For each TCAP entity, this comprehensive resource includes a description, any aliases, their designation status, main language, affiliated media outlets, and ideology. It is accessible here.

LEGAL CLARITY

Tech companies face a lack of clarity on the illegality of far-right terrorist content. There is a broader international consensus and precedent for the designation of Islamist groups, such as Islamic State and al-Qaida, which provides a clear framework within which tech companies can assess their legal obligations relating to the online propaganda of designated groups.

Tech companies assessing their legal obligations to remove far-right content must navigate a myriad of legal and regulatory frameworks that are influenced by overlapping jurisdictional contexts. For example, a National Action propaganda video hosted on a US-based platform will be illegal in the UK (as it is a proscribed terrorist group) but not in the US unless it violates laws relating to incitement to violence.

In Tech Against Terrorism's experience, smaller tech companies often do not fully understand these competing legal and regulatory obligations, and this leads to inconsistency in the moderation of far-right content. Alt-right platforms, on which a significant proportion of far-right terrorist content has been identified, often base their moderation rules strictly on local laws and will only remove content when there is a legal obligation to do so.

Recommendations:

- Tech Against Terrorism recommends that tech companies prohibit both terrorism and violent extremism in their terms of service / Community Guidelines. A prohibition of terrorism should include content which encourages, supports, glorifies and/or promotes terrorism, terrorist organisations, terrorist attacks and attackers. The prohibition of violent extremism allows platforms to more effectively moderate violent far-right content that has not been produced by designated far-right groups.

- Tech Against Terrorism recommends consulting national and supranational designation lists as a guide to the far-right entities that have been designated as terrorist through a stringent legal process. While the legality of online content produced by these groups is contingent on jurisdictional considerations, designation lists are widely accepted as a strong legal basis for content removal. You can find a list of designated far-right entities here (and in *Fig. 2*).

- Beyond designation lists, tech companies should consider other legal instruments to guide their moderation of far-right TVE content. Tech Against Terrorism recommends focusing on content produced by far-right terrorist attack perpetrators, such as the manifesto and livestream produced by the 2019 Christchurch attacker. There are various legislative frameworks that cover this type of far-right content including New Zealand's Classification Office, Australia's regulation of abhorrent material, the European Union's (EU) Terrorist Content Online (TCO) Regulation, and the UK's Online Safety Act (OSA). Explore how we legally classify this type of content here.

- OFCOM has released its Online Safety Guidance for Judgement for Illegal Content to support online service providers' assessment of illegal content under the Online Safety Act. Although still in the consultation phase, this includes guidance on what might be considered far-right terrorist content and provides useful examples such as the Norway manifesto and James Mason's *Siege*.[33]

- All tech companies that offer services globally are likely to have regulatory obligations that exist outside the country in which they are based. Tech Against Terrorism provides handbooks for understanding this complex regulatory landscape with its Online Regulation Series. Companies that operate in European markets must now comply with the EU TCO, which includes obligations for the removal of far-right terrorist content. Further information

[33] The book Siege is a collection of editorials from a monthly newsletter, produced by James Mason from 1980-1986, which promotes neo-Nazism and lone-wolf terrorism. Since the original was published in 1992, there have been three more editions (2003, 2017, 2018) with new prefaces, appendices and added images.

on these obligations can be found here and on the Tech Against Terrorism Europe website.

TRANSPARENCY

Our analysis of tech platform responses to far-right terrorist content has highlighted a gap in our understanding of how platforms enforce their own policies on this type of content. This is due to a lack of clarity in defining what specific TVE content is prohibited in the platforms' public policies and a lack of transparency on the enforcement of internal guidelines and policies.

Recommendations:

- Tech Against Terrorism recommends that tech companies clearly explain what is prohibited on their services. These policies should be clear and easily understandable for users.
- We recommend that a prohibition of terrorism should include:
  - What the platform considers to be terrorism / a terrorist entity (if no definition is provided, platforms should at least reference the designation lists they consult).
  - What the platform considers to be terrorist content, specific to the type of content that can be found on its services such as usernames, a live-streamed video, a video, a user's avatar etc.
  - Illustrative examples for each type of content.
  - The available enforcement actions for violations of prohibition of terrorism, including any warning system or immediate removal enforcement policies.
- Tech Against Terrorism recommends informing users why their content or account was actioned, and what specific policy was violated. For transparency purposes, this information can also be provided on the webpage from which the content was removed. A redress mechanism should be available for users to challenge moderation decisions.
- Tech Against Terrorism recommends that tech platforms produce a transparency report about their moderation enforcement actions. Regularly (annually or bi-annually) publishing a transparency report would allow the platform to significantly increase its transparency and accountability, in addition to alleviating user concerns about privacy. This would also demonstrate the platform's efforts to counter terrorist and violent extremist use of its services.
- For more guidance, find this dedicated resource on our Knowledge Sharing Platform (KSP).
- The EU Digital Services Act details requirements for providers of intermediary services to publish transparency reports on the content moderation actions they have taken in the relevant period at regular annual intervals.

4. FORWARD LOOK

EXISTING RESPONSES TO FAR-RIGHT TERRORIST EXPLOITATION ONLINE

The pervasive nature of social media and the broad audience it offers provides a range of harmful actors, including far-right terrorists and violent extremists, with an unprecedented platform to amplify their messages, recruit followers, and coordinate real-world acts of violence. The challenge for online platforms and authorities is to strike the delicate balance between freedom of speech and preventing the spread of dangerous ideologies that may incite violence. Stakeholders ranging from government, digital industry, civil society, and academia have collaborated and coordinated on initiatives to minimise the risks posed by terrorists and violent extremists, including those inspired by far-right ideologies and grievances.

CHRISTCHURCH CALL TO ACTION

Efforts to curb the influence of far-right violent extremists online must involve collaboration between digital industry, governments, civil society, and academia. A notable example of collective action to reduce the spread of terrorist and violent extremist content online, including far-right-inspired content, is the Christchurch Call to Action (Christchurch Call). The Christchurch Call is a community of over 130 governments, online service providers, and civil society organisations, including Tech Against Terrorism, that are working to advance the collective commitments outlined in the Christchurch Call. This is focused on, but not limited to, understanding the use of algorithms and developing algorithmic interventions, developing and implementing a crisis response mechanism for real-world acts of terrorism or violent extremism, and increasing the overall transparency of relevant actors. Tech Against Terrorism works closely with the Christchurch Call community to advance this work through its participation in the Christchurch Call Advisory Network and, in this capacity, attending the Christchurch Call Leaders' Summit in 2023.

GOVERNMENT RESPONSES

In recent years, several governments have taken significant steps to respond to the increasing threat posed by far-right terrorist exploitation of online platforms. The expansion of terrorist designation lists to include more far-right terrorist entities has been particularly impactful. The designation process is the international system by which governments can classify either a group or an individual as a 'terrorist' entity. The designation of far-right terrorist entities supports online platforms in moderating their services by reducing the complexity of defining terrorist entities and providing a legal justification for the removal of content. Canada and the United Kingdom have, to date, designated the most far-right groups as terrorist.

TCAP EXPANSION

The TCAP Inclusion Policy has steadily expanded to adapt to the evolving far-right terrorist threat. As of March 2024, we alert propaganda content produced by 14 designated far-right entities, and content produced by 6 different far-right terrorist attack perpetrators. We also alert promotional content that glorifies far-right entities within scope of the TCAP, which is currently limited to gamified versions of attack livestreams.

Within the existing framework of the Tiered System, Tech Against Terrorism will be further expanding the far-right terrorist content that we alert to tech companies in the future:

CRISIS CONTENT (TIER 2)

Tech Against Terrorism considers 'crisis content' to be content directly and verifiably produced by terrorist attack perpetrator(s) or their associates which depicts, justifies, or amplifies their actions or motivations. Eligible content can include a manifesto, livestream, or other relevant material such as a video or statement.

We are assessing historical crisis events involving content produced by far-right attackers to identify priority content for future inclusion. Attacker-produced content under consideration includes content relating to the following incidents: **Hanau 2020, El Paso 2019, Charleston 2015, Poway 2019, and Isla Vista 2014.**

DESIGNATED ENTITIES (TIER 3)

There are limitations to relying exclusively on designation for countering the far-right TVE threat online. As our analysis reveals, many of the designated far-right groups in scope of the TCAP have produced limited publicly accessible propaganda online. The groups for which we have identified least official propaganda are Combat 18, Russian Imperial Movement (RIM), National Socialist Anti-Capitalist Action (NS131), Scottish Dawn, System Resistance Network, and the Proud Boys. This suggests designation has been effective at undermining the overt online presence of these groups, either dismantling the group or forcing them into private online spaces.

Despite its limitations, designation lists remain a useful reference and important legal tool for guiding TCAP Inclusion. It allows us to capture the most dangerous far-right terrorist actors and their violent brands such as AWD, The Base, and National Action.

To ensure designation is utilised most effectively to counter the far-right threat, Tech Against Terrorism will:

ENCOURAGE THE TIMELY DESIGNATION OF FAR-RIGHT GROUPS

- Tech Against Terrorism has consistently argued that governments should consider designating more far-right terrorist groups to more accurately reflect and respond to the danger posed by the transnational far right. For more detailed recommendations, refer to our report *'Who Designated Terrorism? The Need for Legal Clarity to Moderate Terrorist Content Online'*.

- When assessing whether far-right groups meet the legal threshold for designation, governments should ensure they are closely monitoring and considering the online activity of the group and its affiliated networks, factoring in the prevalence of content that encourages terrorism or incites violence.

- Violent far-right groups often operate online as decentralised networks with loosely affiliated individuals carrying out violence. One such group is the Order of the Nine Angles (09A), a UK-based Satanic neo-Nazi group; there have been widespread calls for its proscription.[34] Designating groups of this nature provides a strong legal basis for service providers to remove their online propaganda and undermines the group's ability to radicalise internet users towards violence. The widespread accessibility of 09A content, including on mainstream online marketplaces and in comparison with propaganda from designated groups, further emphasises this point.

- Tech Against Terrorism will continue to monitor non-designated far-right violent extremist groups through open-source intelligence and share this information with governments to encourage the timely designation of these entities where appropriate and with platforms to inform their threat assessments and policies.

ENSURE OFFSHOOTS OF EXISTING DESIGNATED GROUPS ARE QUICKLY AND ACCURATELY IDENTIFIED AS ALIASES

- Designation lists often include named aliases for existing groups to ensure they cannot circumvent legal restrictions.

- Designating authorities such as the UK should consider designating offshoots of National Socialist Order as aliases. This includes the National Socialist Resistance Front (NSRF) and National Socialist Order of the Nine Angles (NSO9A).

PROMOTIONAL CONTENT (TIER 4)

Inspirational material is a subset of Tier 4 and defined as content that "explicitly encourages, glorifies and/or incites a terrorist act or praises the perpetrator(s) of that act, given the entity (individual or organisation) is included within scope of the TCAP."[35] A key aim of this category is to counter the sanctification of terrorist attack perpetrators.[36]

We maintain an internal list of inspirational material, which we began alerting through Tier 4 of the TCAP as of April 2024. Although the full list will not be public due to concerns around publicising this terrorist material, it includes:

- Terrorgram publications and videos
- Militant accelerationist content
- Content affiliated with satanic occultism

[34] Source: https://hopenothate.org.uk/wp-content/uploads/2020/02/state-of-hate-2020-final.pdf

[35] Source: https://www.terrorismanalytics.org/policies/promotional

[36] Source: https://gnet-research.org/2023/04/27/the-lineage-of-violence-saints-culture-and-militant-accelerationist-terrorism/

THE UTILITY OF TECHNICAL SOLUTIONS

HASH-MATCHING

The utility of hash-matching for content moderation is well-known and widely evidenced for a range of online harms. The TCAP Archive, which is currently under development, will be the first archive of terrorist content open to all platforms which provides for hash-match detection of both content and hashes.

The nature of much of the far-right content observed in TCAP submissions highlights some of the identified difficulties with hash-matching technology. With the most visually distressing element of TCAP far-right alerts encompassing livestreamed videos, often of considerable length, we are focused on enhancing perceptual hashing capabilities to be able to identify the many variants of such livestreams we observe.

We also aim to develop hashing capabilities to enable the matching of ‘chunks’ of text within a given document rather than requiring the full file. We have collected archives of content for the range of included far-right entities under TCAP, in line with our goal to ensure far-right content is considered as seriously and understood as well as violent Islamist material.

GENERATIVE AI

Recognising some of the limitations regarding hashing technology, the TCAP is currently building a large-language model (LLM) to assist in more sophisticated detection and classification of terrorist content.

Trained on the vast corpus of terrorist material we will be hosting in our Archive, the LLM will be attuned to relevant logos, symbols, depictions, and themes present in far-right content.

As we use open-source intelligence to identify more and more examples of such propaganda, the model will be refined and is likely to prove a highly useful addition for moderating new content that has never been hashed previously – especially that which is created by generative AI tools.

ANNEX

TCAP INCLUSION POLICY FOR FAR-RIGHT TERRORIST CONTENT

DEFINING FAR-RIGHT TERRORIST CONTENT

The purpose of this report has not been to engage with definitional debates around what constitutes the far-right. However, given that we are referring to 'far-right terrorist content' throughout the report it is important to set out how we define the scope of inclusion. Given the lack of international consensus as to what far-right terrorism is, Tech Against Terrorism's approach is to defer to terrorism legislation at the national level, primarily referring to terrorism designation lists. Beyond terrorist designation lists, and to account for the post-organisational nature of the far-right threat, we also include within the TCAP promotional terrorist material that directly supports designated terrorist organisations or glorifies terrorism with close reference to the European Union Terrorist Content Online (TCO) legislation and the United Kingdom Terrorism Act (TACT). We outline this approach in more detail below:

TCAP INCLUSION POLICY FOR FAR-RIGHT CONTENT

Below, we outline the scope of far-right terrorist content included within this report based on the TCAP's Tiered Inclusion Policy.[37]

TIER 2 – CRISIS

DEFINITIONS

Event: An act of real-world violence that is carried out by a non-state actor with the intent to endanger, cause death or serious bodily harm to a person[s] and is motivated by **ideological, political, or religious goals**.

Crisis content: Content directly and verifiably produced by a terrorist attack perpetrator or perpetrators during an **event** depicting, justifying or amplifying their actions or motivations and/or inciting others to commit acts of violence. In practice, this means a manifesto, livestream or other relevant material such as a video or statement produced by the attack perpetrator(s) or their associates.

For far-right attacker-produced content to be in scope, it must as a minimum meet the above criteria of being violent extremist material produced by the perpetrator of a terrorism-related event.

Given the wide scope of content this could potentially include, we have prioritised the inclusion of material that has been classified by the New Zealand Classification Office as 'objectionable.' Classified material is illegal in New Zealand and this mechanism provides an additional legal basis for our inclusion. You can find the material, the related incident, and when it was added to the TCAP below:

[37] You can find our full Inclusion Policy on the TCAP website here.

FAR-RIGHT TERRORIST INCIDENTS IN SCOPE OF TCAP INCLUSION POLICY

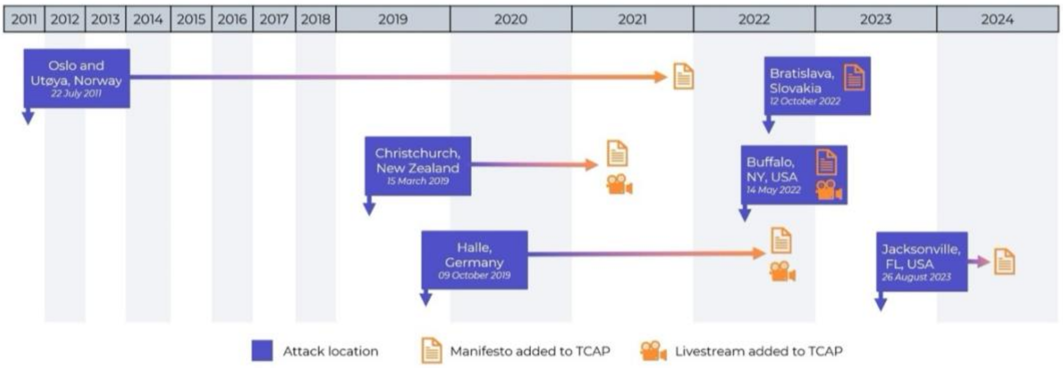
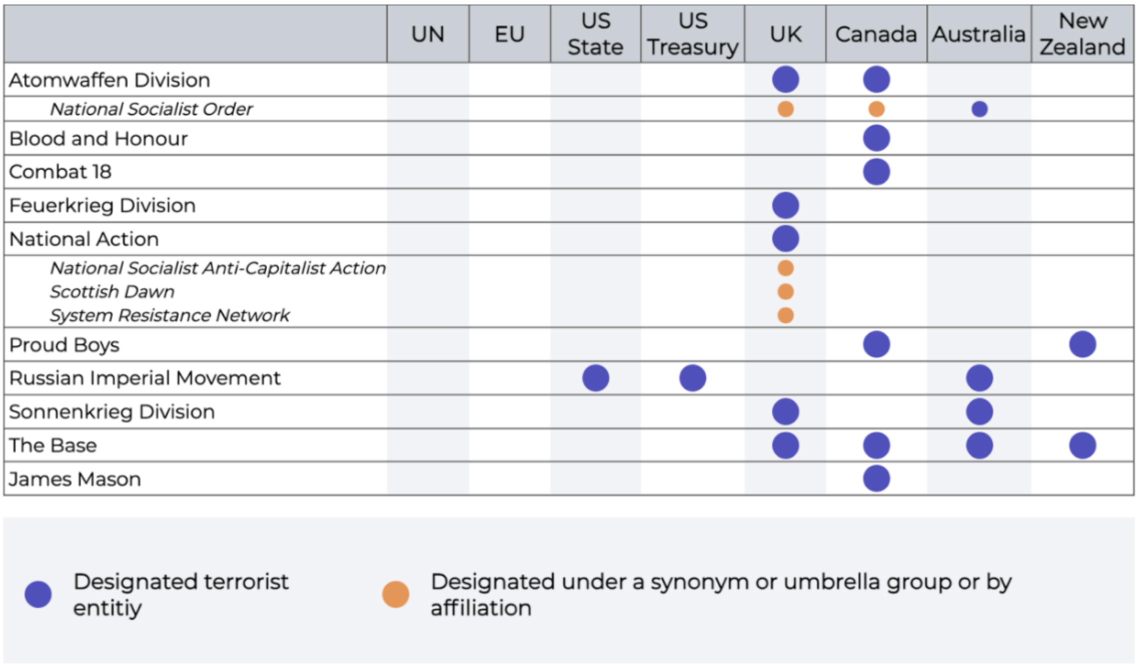

2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
Oslo and Utøya, Norway 22 July 2011
Bratislava, Slovakia 12 October 2022
Christchurch, New Zealand 15 March 2019
Buffalo, NY, USA 14 May 2022
Halle, Germany 09 October 2019
Jacksonville, FL, USA 26 August 2023
Attack location
Manifesto added to TCAP
Livestream added to TCAP

Figure 33: Far-Right Terrorist Incidents in Scope of TCAP Inclusion Policy.

TIER 3 – DESIGNATION

To be in scope, a far-right entity needs to be designated as terrorist by at least one of the following authorities: the European Union, the United Nations, the United States, Canada, New Zealand, Australia, and the United Kingdom. As such, the TCAP currently issues alerts for official propaganda produced by the following 14 far-right terrorist entities:

FAR-RIGHT DESIGNATED TERRORIST ENTITIES INCLUDED IN TCAP



	UN	EU	US State	US Treasury	UK	Canada	Australia	New Zealand
Atomwaffen Division					●	●		
National Socialist Order					🟠	🟠	●	
Blood and Honour						●		
Combat 18						●		
Feuerkrieg Division					●			
National Action					●			
National Socialist Anti-Capitalist Action *Scottish Dawn* *System Resistance Network*					🟠 🟠 🟠			
Proud Boys						●		●
Russian Imperial Movement			●	●			●	
Sonnenkrieg Division					●		●	
The Base					●	●	●	●
James Mason						●		

● Designated terrorist entitiy

🟠 Designated under a synonym or umbrella group or by affiliation

Figure 34: Far-Right Designated Terrorist Entities for TCAP Inclusion Policy.

TIER 4 – PROMOTIONAL

Introduced in July 2023, Tier 4 expanded the scope of the TCAP to include material that directly supports designated terrorist organisations, glorifies terrorists or terrorist acts, or provides instructions for the purpose of terrorism.

INSPIRATIONAL MATERIAL

Inspirational material is a sub-tier of Tier 4 that includes "*content that explicitly encourages, glorifies and/or incites a terrorist act or praises the perpetrator(s) of that act, given the entity (individual or organisation) is included within scope of the TCAP.*"

Within the data collection period which ended in October 2023, the only far-right content in scope of this Tier comprised gamified versions of the livestreams of the following attacks:

- Christchurch, NZ
- Buffalo, NY
- Halle, Germany

OTHER TCAP POLICIES & SERVICES

INCIDENT RESPONSE POLICY

Increasingly, far-right terrorist actors are exploiting online services to maximise the reach of their message and exposure to their violent acts, such as through disseminating a livestream or manifesto. A large proportion of the livestreamed attacks and manifestos have been produced by far-right terrorists carrying out 'lone-wolf' attacks. Recognising this threat, Tech Against Terrorism has developed an Incident Response Policy to minimise the impact of terrorist attacks by inhibiting the spread and potential virality of attacker-produced content online.

The mechanism aims to disrupt hostile responses of online terrorist and violent extremist networks which often function as incubators of hateful and violent grievances. The objective is to facilitate rapid and targeted disruption of terrorist content on smaller platforms. Furthermore, the policy directly meets and advances the Christchurch Call to Action commitments, including the commitment to support smaller platforms, as they build capacity to reduce the spread of terrorist and violent extremist content online, and the commitment to develop processes to respond rapidly and effectively to the dissemination of terrorist or violent extremist content following a terrorist event. Tech Against Terrorism strongly believes the Incident Response Policy will support a reduction in the spread of far-right terrorist content and activity online.

TCAP ARCHIVE

Tech Against Terrorism is developing the Terrorist Content Analytics Platform (TCAP) Archive to support online platforms in training content moderators to identify and analyse terrorist content on their services. The archive is comprised of the historical content collected and alerted through the TCAP, including designated far-right terrorist content. The use of the archive will support online platforms to undertake more efficient and accurate detection and removal of far-right terrorist content. Specifically, it will provide a trove of far-right terrorist content that can be used by online platforms to train content moderators to better identify this type of content on their platforms.

In addition, the archive will include the capability for online platforms to submit hashes of suspected terrorist content and receive a positive or negative match against the archived content. This functionality will support smaller platforms in particular in moderating terrorist content online by providing confirmation of verified terrorist content. Tech Against Terrorism expects the TCAP Archive to be operationalised in 2024.

METHODOLOGY

DATA COLLECTION

The analysis in this report is based on data relating to terrorist content online collected by the TCAP between 20 February 2021 and 1 November 2023. This dataset includes 2,966 URLs of terrorist content submitted to the TCAP (TCAP Submissions), including 2,348 of those sent as alerts to 55 different tech companies (TCAP Alerts). This dataset also includes the removal rate of terrorist content alerted to tech companies – namely, whether a platform removes terrorist content after it has been alerted to them via the TCAP. The status of URLs used for the report was last checked on 31 October 2023.

TCAP SUBMISSIONS AND ALERTS[38]

Submissions: Tech Against Terrorism's open-source intelligence team monitor and identify terrorist content daily. Each piece of content is verified against the TCAP Inclusion Policy and attributed to the corresponding terrorist organisation. Once content has been verified and classified, it is uploaded to the TCAP (submission).

Alerts: Once content is submitted, the TCAP emails the platform in question with the link to where the content can be found, the associated terrorist entity, and a warning for content that is graphic in nature or contains personally identifiable information. TCAP alerts are made on an advisory basis, meaning it is the platform's decision to proceed with content moderation.

[38] For further information on the methodology of the TCAP, please see Terrorist Content Analytics Platform, How it works.

Terrorist Content
Analytics Platform

Powered by

tech
against
terrorism